MAR 1 6 1922

MAR 1 1922

The Journal of Educational Psychology

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching.

BOARD OF EDITORS:

HAROLD ORDWAY RUGG, *Chairman.*
Lincoln School of Teachers College.
Teachers College, Columbia University.

JAMES CARLETON BELL,
Brooklyn Training School for Teachers.

FRANK NUGENT FREEMAN,
University of Chicago.

ARTHUR IRVING GATES,
Teachers College, Columbia University.

VIVIAN ALLEN CHARLES HENMON,
University of Wisconsin.

RUDOLF PINTNER,
Teachers College, Columbia University.

BEARDSLEY RUML,
Carnegie Corporation, New York City.

LEWIS MADISON TERMAN,
Leland Stanford University.

EDWARD LEE THORNDIKE,
Teachers College, Columbia University.

LAURA ZIRBES, *Assistant Editor.*
Lincoln School of Teachers College.

VOL. XIII.
No. 3.

MARCH, 1922

$4.00 a year.
60c. per copy.

CONTENTS

THE CORRELATIONS OF ACHIEVEMENT IN SCHOOL SUBJECTS WITH INTELLIGENCE TESTS AND OTHER VARIABLES. *Arthur I. Gates* 129

INTELLIGENCE IRREGULARITY AS MEASURED BY SCATTERING IN THE BINET SCALE. *J. E. W. Wallin* 140

A STUDY OF HIGH SCHOOL SPELLING MATERIAL. *John A. Lester* 152

A COMPARATIVE STUDY OF A BORDER LINE DEFECTIVE AND A NORMAL CHILD OF THE SAME MENTAL AGE. *Laura Remer* 160

JUDGING HANDWRITING. *Willard W. Beatty* 170

THE TORONTO MEETING OF THE SECTIONS ON PSYCHOLOGY AND EDUCATION OF THE AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE. *Frank N. Freeman* 173

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES 177

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION 180

Published Monthly Except June to August by
WARWICK and YORK, Inc.,
York, Pa. *Baltimore, Md.*

Entered as Second Class matter November 15, 1921, at the Post Office at York, Pennsylvania, under the Act of March 3, 1879.

The Journal of Educational Psychology

is published monthly, by Warwick & York, Inc., Baltimore, Md., and York, Pa. The issues for 1922 make Volume XIII. Title page and index are bound in December number.

Manuscripts for publication, books or other materials for review, and news items should be addressed to Harold Ordway Rugg, Chairman of the Editorial Board, 59 Edgecliffe Terrace, Yonkers, New York.

The price of the Journal is $4.00 a year, payable in advance. Foreign postage is 30 cents extra. Single issues are 60 cents each, and less than a full year's subscription is charged at single issue rate for every month ordered.

Subscribers should notify the publishers of change in address at least four weeks before publication of the issue with which change is to take effect. No claim for non-receipt of an issue will be entertained unless made within two weeks after receipt of the next succeeding issue.

Warwick & York, Inc.,
Baltimore, Md.

TEACHERS' HEADQUARTERS

European
Plan

Entirely
Fire-Proof

Centrally
Located

EDWARD DAVIS
Manager

HOTEL RENNERT
BALTIMORE, MARYLAND

University of Virginia Summer Quarter

First Term, June 19–July 29
Second Term, July 31–September 2

Courses for Elementary Teachers
Courses for High School Teachers
Courses for College Credit

The Summer Quarter is an integral part of the University Year, the courses being the same in character and credit value as in the other quarters of the year.

Degrees are conferred upon men and women for summer work.

The Master's Degree may be obtained in three Summer Quarters.

It offers opportunities unexcelled in the South and makes a strong appeal to teachers seeking broader scholarship and training, and wider social contacts, and to college students desiring to complete degree requirements.

Attendance last Quarter, 2429 from twenty-nine states and foreign countries.

The most beautiful and unique campus in America.

Accommodations at reasonable rates. Tuition for non-Virginia students $15.00 per term.

Entertainments, Music Festival, excursions.

For illustrated folder and full announcement, write to

CHAS. G. MAPHIS, Dean,
University, Virginia.

The University of Wisconsin Summer Session

June 26 to August 4 (Law School, June 19 to August 25)

One Fee, $20, for ALL Courses (Except Law $35)

Courses in all Colleges leading to the Bachelor's and Higher Degrees

Special courses in the psychology of learning, experimental studies in learning, statistical methods, individual differences and the measurement of intelligence, the psychology and education of exceptional children, school surveys and tests, the psychology of high school subjects, mental development, and social development.

Research Facilities *Lakeside Advantages*

For further information address: Director Summer Session, Madison, Wis.

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Vol. XIII March, 1922 No. 3

THE CORRELATIONS OF ACHIEVEMENT IN SCHOOL SUBJECTS WITH INTELLIGENCE TESTS AND OTHER VARIABLES[1]

ARTHUR I. GATES

Teachers College, Columbia University

Many tests of intelligence are now available. They are classified in several ways; most commonly as verbal and non-verbal, although many combine both varieties of exercises. Whether the extremes have equal predictive value is but imperfectly known. In one sense they have equal predictive value, since perhaps every function is closely associated with certain other functions. The question is: Just what functions are correlated highly with verbal, non-verbal or any other abilities?

It has often been demonstrated that certain tests have a high reliability; that is, they measure consistently whatever it is that they do measure. The Stanford-Binet measures something very well but of just what abilities it is a valid measure or for just what abilities it gives a valid prediction, is far from perfectly known.

Usually the test has been developed to yield a predictive measure of some particular ability or composite of abilities. Often the tests reach a high state of mechanical perfection, while the criterion—the abilities to be predicted—remains unanalyzed and but imperfectly measurable. In the field of educational prediction, the crucial problem is the perfection of criteria. We secure little or no information about a test by computing correlations with teachers' estimates, which may be good or bad—mediocre, in all likelihood—but just which is never known. If the actual correlation of the Stanford-Binet and real "general intelligence" were approximated unity, we would not dis-

[1] A contribution of the Department of Educational Research of the Scarborough School, Scarborough, N. Y.

cover it by comparisons with the conventional school grades. We need better criteria.

In this study, we have taken as our problem the prediction of achievement in the fundamental school subjects. A special effort has been made to secure measures of achievement which are as valid and reliable as facilities would permit. Each school subject has been measured by a large number of standard tests: In some cases, *e.g.*, reading in certain grades, nearly 10 hours work is represented. The reliability of the composites is high; the coefficients of reliability (self correlations) being in all cases between 0.92 and 0.97.

We cannot be so sure of the validity of the composites. The series of arithmetic tests, even if self consistent, may be an inadequate measure of "general arithmetical ability." Without an ultimate criterion outside the tests themselves, the validity of the tests cannot be determined. The different tests should be weighted by use of the regression coefficients, but with no ultimate criterion the perfect weights cannot be determined. Mere accumulation of inadequate tests or of tests which measure well only a common fraction of the "general" ability will not, of course, insure a valid criterion. The best one can do is to select the tests which are most reliable, which test as diverse elements within the general function as possible, and weight them arbitrarily.

Even with reliable and valid tests, properly weighted, as a criterion, perfection is not necessarily attained. The effects of school and home emphasis on particular subjects may operate to reduce the correlation with the predictive variables. Arithmetic may be so stressed that the pupils have attained nearly maximal performance ability but spelling may be so neglected that few if any have it, even approximately. In a second school, conditions may be the reverse. The predictive value of the intelligence tests will almost certainly differ in the two schools. The ideal would be a situation in which each pupil had reached his approximate limit of improvement which is then measured by valid and consistent instruments. This situation will probably not be obtained except under strictly experimental conditions.

The Tests Employed

1. *Tests of Achievement:*

The measures for grades I and II are very much less extensive and valid than those in the other grades. The tests given were:

(*A*) For grade I. Indiana Reading vocabulary, Haggerty, Sigmad I, and 50 words from Ayres Spelling list.

(*B*) For grade II. Indiana Reading vocabulary, Haggerty, Sigma I, Thorndike-McCall Reading, 50 words from Ayres Spelling and the Indiana Composite Scale of Attainment, No. 2, including tests of reading, word knowledge, arithmetic, and spelling.

The following tests were given in grades III to VIII inclusive:

(*A*) *Composite of Comprehension in Reading.*—The scores of the Burgess, Courtis, Monroe, Thorndike, Thorndike-McCall, and Woodworth-Wells Directions tests were combined. The scores for certain tests were the means of several tests in which different forms or editions of the test were used. In grades IV and VI the Thorndike-McCall was given five times. The SD's were such that the several tests were given approximately equal weights. This is true of the following composites.

(*B*) *Rate of Reading.*—A composite of the Burgess, Courtis Rate, Brown Rate and Monroe Rate, all given at least twice and weighted equally.

(*C*) *Arithmetic.*—A composite of the Woody 30-minute test for each of the four operations, the Monroe Diagnostic, 11 to 21 specific tests, ranging from easy integer combinations to fractions and decimals, and the Monroe Reasoning Test.

(*D*) *Spelling.*—Four tests making a total of 186 words of varied difficulty from the Ayres-Buckingham list.

(*E*) *Writing.*—Two sets of specimens graded for quality by the Thorndike scale. Speed and quality were combined by multiplying the number of letters per minute by the quality score.

(*F*) *Composite of General Achievement.*—A combination of reading comprehension, reading rate, arithmetic and spelling. The SD's were such that each received approximately the same weight.

2. *Other Criteria Used:*

(*A*) *Teachers' Estimates of School Attitudes.*—A rating scale requesting judgments on a composite of industry, interest, attention, etc. was filled out for each pupil by from 5 to 9 teachers or supervisors, independently. Only the ratings for grades IV to VIII are used in this study.

(*B*) *Chronological age.*

3. *The Predictive Measures:*

(*A*) *Stanford-Binet Mental Age*, grades I to VI, inclusive.

(*B*) *Group Tests of Intelligence.* (1) Dearborn, Examinations 1, 2,

3, and Total, (2) Dearborn, Examinations 4, 5, and Total, (3) Haggerty, Delta 1, (4) Haggerty, Delta 2, (5) Holley's Picture Completion, (6) Holley's Sentence Vocabulary, (7) Illinois Examination, (8) Kingsbury Primary, Form A, (9) Meyers Mental Measure, (10) National Intelligence, Form A, B, and Total, (11) Otis, Primary, Form A, (12) Otis, Advanced, (13) Pressey Primer, (14) Terman Group, (15) Thorndike-McCall Reading Test.

These tests were given as follows: To grades I and II, Nos. 1, 3, 5, 8, 9, 11 and 13; grade III, Nos. 1, 3, 6, 7, 9, 10, 11 and 15; grades IV 2, 4, 5, 6, 7, 9, 10, 11 and 15; grade V to VIII inclusive, 2, 4, 5, 6, 7, 9, 10, 12, and 15. The Terman Group Test, grades VII and VIII.

The Subjects.—The subjects were the pupils of grades I to VIII inclusive, in the Scarborough School, at Scarborough, N. Y. These are select groups; the mean Stanford-Binet Intelligence Quotients are about 117. There were about 20 pupils to the grade.

Statistical Methods Employed.—All coefficients of correlation were computed by the Pearson Product-Moment formula. The tables give the correlations for each variable for each grade separately. The mathematical estimates of the probable error of the correlations which are based on the number of cases are not included, but may be readily estimated. The variability of the separate grade (or test) correlations from the mean of the grades (or tests) gives the best notion of their reliability. The mean is used as a measure of central tendency of the r's and the SD's give the variability.

Because of the restriction in the range of abilities in our groups (children of like abilities being selected for a grade) the correlations are certainly lower than they would be for an unselected group. There is no technique for correcting the attenuation due to restriction in range, which could be applied to our data.

In so far as the measures used fail to measure a subject's abilities perfectly, the correlations will again be attenuated. Many of our measures have a high degree of reliability because of the thoroughness of the testing and since our main purpose is make comparisons of one variable with another, within the data for our own groups, we have made no correction for this type of attenuation. It will, of course, be understood that the absolute amounts of our coefficients may not be compared with those obtained from other groups unless the range should happen to be the same.

In many cases, the inter-correlations are so many and varied that they cannot be readily interpreted by inspection. It has been advis-

able to employ, in certain cases, the technique of multiple and partial correlations, when the most perfect weight of each variable has been determined by the regression equations.

There are now several methods by which these data may be obtained. The most recent and excellent procedure in print is that devised by Dr. T. L. Kelly.[1] Dr. Herbert A. Toops has more recently devised (but not yet printed) a different and more expeditious method of obtaining multiple correlations which will be explained and defended by him in due time. All of the partial correlations presented in this article were computed by Dr. Kelley's method and the multiple correlations by Dr. Toops' method. In certain cases multiple correlations and partial correlations of the first order were computed by both methods and found to agree.

Part I. Results of Tests in Grades I and II

Table I gives the inter-correlations of the group tests, and the correlations of the group tests with Stanford Mental Age and with the composite of tests of achievement for each grade, together with the average results for the two grades.

1. *Correlations with the Composite of Achievement.*—Averaging the results of all tests for the two grades the following correlations are found:

1. Achievement with MA	0.40
2. Achievement with Group Tests[2]	0.27
3. Achievement with Chronological Age	0.15
4. MA with Group Tests	0.43
5. MA with Chronological Age	0.27
6. Group Tests with Chronological Age	0.47

The inter-correlations are too complicated to permit ready interpretation by inspection. While MA yields a higher correlation with achievement than any single test, group tests and age also show positive correlations. Do the latter criteria add anything not included in the MA? If so what and how much?

---

[1] Kelly, T. L.: Chart to Facilitate the Calculation of Partial Coefficients of Correlation and the Regression Equations. Stanford University Publication, School of Education, *Monograph* No. 1, 1921.

[2] This is not the correlation with a composite of group tests; it is the correlation of the average (single) group test when the r's for the two grades are averaged.

TABLE I.—CORRELATIONS FOR GRADES I AND II

Grade	1 Dearborn 1	2 Dearborn 2	3 Dearborn 3	4 Dearborn total	5 Otis Primary	6 Kingsbury	7 Haggerty Delta 2	8 Pressey Primer	9 Holley Picture	10 Myers	11 Chronological Age	12 Stanford MA	13 Composite Achievement	14 Mean Group Tests	15 Mean all criteria	16 Time (minutes)
Correlations of Dearborn, Examination 1 with																
I	....	0.84	0.67	0.91	0.82	0.53	0.55	0.34	0.45	0.62	0.48	0.53	0.42			
II	....	0.43	0.75	0.82	0.56	0.14	−0.13	0.43	0.18	0.51	0.42	0.54	0.10			
Mean	....	0.64	0.71	0.87	0.69	0.34	0.21	0.39	0.31	0.57	0.45	0.54	0.26	0.48	0.47	40
Correlations of Dearborn, Examination 2, with																
I	0.84	....	0.69	0.95	0.71	0.46	0.53	0.21	0.32	0.47	0.44	0.53	0.53			
II	0.43	....	0.36	0.61	0.56	0.25	0.49	0.31	0.25	0.41	0.40	0.32	0.38			
Mean	0.64	....	0.53	0.78	0.64	0.36	0.51	0.26	0.29	0.44	0.42	0.43	0.46	0.46	0.46	40
Correlations of Dearborn, Examination 3, with																
I	0.67	0.69	....	0.81	0.70	0.52	0.50	0.51	0.60	0.73	0.43	0.52	0.42			
II	0.75	0.36	....	0.81	0.67	0.18	0.17	0.51	0.20	0.49	0.49	0.58	0.12			
Mean	0.71	0.53	....	0.81	0.69	0.35	0.34	0.51	0.40	0.61	0.46	0.55	0.27	0.52	0.50	40
Correlations of Dearborn, Total, with																
I	0.91	0.95	0.81	....	0.76	0.52	0.63	0.35	0.49	0.63	0.53	0.58	0.48			
II	0.82	0.61	0.81	....	0.69	0.30	0.24	0.59	0.25	0.68	0.45	0.56	0.32			
Mean	0.87	0.78	0.81	....	0.73	0.41	0.44	0.47	0.37	0.66	0.49	0.57	0.40	0.51	0.51	120
Correlations of Otis Primary with																
I	0.82	0.71	0.70	0.76	....	0.45	0.53	0.36	0.30	0.68	0.40	0.46	0.38			
II	0.56	0.56	0.67	0.69	....	0.58	0.61	0.54	0.38	0.75	0.38	0.55	0.20			
Mean	0.69	0.64	0.69	0.73	....	0.52	0.57	0.45	0.34	0.72	0.39	0.51	0.29	0.60	0.56	22
Correlations of Kingsbury with																
I	0.53	0.46	0.52	0.52	0.45	....	0.24	0.40	0.80	0.47	0.54	0.56	0.10			
II	0.14	0.25	0.18	0.30	0.58	....	0.26	0.32	0.58	0.51	0.54	0.15	0.08			
Mean	0.34	0.36	0.35	0.41	0.52	....	0.25	0.36	0.69	0.49	0.54	0.36	0.09	0.32	0.38	7
Correlations of Haggerty, Delta I with																
I	0.55	0.53	0.50	0.63	0.53	0.24	....	0.38	0.39	0.45	0.51	0.36	0.21			
II	−0.13	0.49	0.17	0.24	0.61	0.26	....	0.39	−0.03	0.31	0.31	0.44	0.60			
Mean	0.21	0.51	0.34	0.44	0.57	0.25	....	0.39	0.18	0.38	0.41	0.40	0.41	0.36	0.37	12
Correlations of Pressey Primer with																
I	0.34	0.21	0.51	0.35	0.36	0.40	0.38	....	0.54	0.59	0.76	0.24	0.17			
II	0.43	0.31	0.51	0.59	0.54	0.32	0.39	....	0.26	0.38	0.30	0.52	0.26			
Mean	0.39	0.26	0.51	0.47	0.45	0.36	0.39	....	0.40	0.49	0.53	0.38	0.22	0.41	0.39	12
Correlations of Holley Picture Completion with																
I	0.45	0.32	0.66	0.49	0.30	0.80	0.39	0.54	....	0.56	0.38	0.35	0.14			
II	0.18	0.25	0.20	0.25	0.38	0.58	−0.03	0.26	....	0.50	0.40	0.10	0.02			
Mean	0.31	0.29	0.43	0.37	0.34	0.69	0.18	0.40	....	0.53	0.39	0.23	0.08	0.39	0.39	5
Correlations of Myers Mental Measure with																
I	0.62	0.47	0.73	0.63	0.68	0.47	0.45	0.59	0.56	....	0.58	0.16	0.13			
II	0.51	0.41	0.49	0.68	0.75	0.51	0.31	0.38	0.50	....	0.51	0.41	0.25			
Mean	0.57	0.44	0.61	0.66	0.72	0.49	0.38	0.49	0.53	....	0.55	0.29	0.19	0.43	0.49	10
Correlations of Stanford Mental Age with																
I	0.53	0.53	0.52	0.58	0.46	0.56	0.36	0.24	0.35	0.16	0.30	....	0.36			
II	0.54	0.32	0.58	0.56	0.55	0.15	0.44	0.52	0.10	0.41	0.24	....	0.44			
Mean	0.54	0.43	0.55	0.57	0.51	0.36	0.40	0.38	0.23	0.29	0.27	....	0.40			

The questions may be answered by the use of multiple and partial correlations. These have been computed by the formulæ earlier mentioned. The regression equation gives the weights that each variable should be given to predict achievement most perfectly. They are:

7. Weight of MA..... 1.00
8. Weight (β) of Group Tests..... 0.345[1]
9. Weight (β') of Chronological Age..... 0.008[1]

Chronological Age thus appears to add little which is independent of MA and group tests, but the group tests add something to MA which is independent of it.

For practical purposes we wish to know whether the addition of the independent contributions of Group Tests and Chronological Age to MA will increase the correlation with achievement sufficiently to justify the time and expense of administering them. The multiple correlation, each variable weighted perfectly according to its independent addition, will give this information.

10. Simple r, Achievement with MA..... 0.40
11. Multiple r, Achievement with (MA + Group) 0.415
12. Multiple r, Achievement with (MA + Group + CA)..... 0.416

If the Stanford-Binet is given, very little is added to the correlation value by adding a group test or CA, even when they are perfectly weighted. The MA alone is clearly superior to an average Group Test or Chronological Age alone (see 1, 2 and 3 above).

All of the Group Tests (with exception of parts of the Haggerty) are composed of non-verbal materials, so that the results just enumerated cannot be generalized to include verbal material. The relative predictive values of the MA and group tests will be taken up again in connection with the higher grades where verbal as well as non-verbal tests are available.

The validity of the correlations just considered depends upon the validity of the criterion, the composite of achievement. It is less valid than is desirable because of the dearth of adequate educational

[1] β is the weight of group tests independent of the elements already given in the correlation of MA with the criterion; β' is the weight of chronological age independent of the other two. If chronological age were put in second place instead of third it would show a larger weight than it does. It is customary to arrange the variables in the order of the magnitude of their simple correlation with the criterion.

tests for these grades and because difference in pre-school training may considerably affect achievement, particularly in grade I.

2. *The Relation between the Length of a Group Test and Its Correlation with Criteria.*—In Table I the inter-correlations of Group Tests are averaged in column 14. Column 15 gives the mean of all correlations except those with Chronological Age. Column 16 gives the approximate working time for each tests. The group tests are alike in being composed of non-verbal material with the exception of the Haggerty Delta I which contains some verbal material.

Inspection of columns 14, 15 and 16 shows a close relation between the length of the test and the magnitude of the correlations. The range for time, however, is large compared to the range of the correlations. If the ranges are made equal by the use of the rank method, the correlation of time with mean inter-correlations (column 14) is 0.75; of time with correlations for all criteria (column 15) is 0.69. Leaving the ranges as they are, using the Product-moment formula, the correlations become much smaller, *viz.* 0.50 and 0.49. In either case, the longer tests, in general, have a higher predictive value but our data are too few to permit an estimate of the increments in r, which are produced by given additions to the length of the test.

Part II. Correlations in Grade III

Grade III offers an opportunity to test the relative validity of verbal and non-verbal tests for the reason that there are available a number of tests constructed exclusively of each type of material. Our measures of achievement in school subjects are much more extensive and valid in this grade than in the primary grades. The tests used in constructing the composites for the school subjects are listed in the first section of this paper.

Table II gives the detailed results. Tests numbered on lines 1 to 7 inclusive are non-verbal. Line 8 gives the mean inter-correlations for this group. Tests on lines 10–15 inclusive are verbal, wholly or chiefly. Line 16 gives the mean inter-correlations for this group. In computing the means, the correlations of parts of a test with the whole (Dearborn and National) have been omitted since the correlation is made much higher by the correlation of the part with itself in the whole. The total working time for the groups of non-verbal tests is about one-third greater than that required for the group of verbal tests (see column 21).

1. *Correlations of Verbal and Non-verbal Tests with Themselves and Each with the Other.*

From Table II the following means were computed:

Mean inter-correlation of non-verbal tests	0.40
Mean inter-correlation of verbal tests	0.62
Mean correlation, non-verbal with verbal	0.24

The verbal tests quite clearly yield different results from the non-verbal. The verbal tests are more consistent with themselves than the non-verbal but the latter agree among themselves better than with the verbal. Comparative data on particular tests are available in Table II.

2. *The Independent Values of MA Verbal, and Non-verbal Group Tests for Predicting School Achievement.*

The following are the mean simple correlations for this grade:

1.	Achievement with verbal group	0.65
2.	Achievement with MA	0.47
3.	Achievement with non-verbal	0.22
4.	Verbal group with MA	0.47
5.	Verbal group with non-verbal	0.24
6.	MA with non-verbal group	0.16

The verbal group test appears to be the best single predictive measure, followed closely by MA. The correlation of MA and the verbal group is the same as MA with achievement. The other correlations are lower.

Does MA or non-verbal group really add anything unique to the verbal group for predicting achievement?

By use of the regression equation the following weights were found.

7.	Weight of verbal group	1.000
8.	Weight (β) of MA	0.499
9.	Weight (β') of non-verbal	0.097

Non-verbal adds very little not already accounted for by MA and the group tests, while MA shows a fair contribution, independent of the verbal tests.

The following multiple correlations show to what extent the addition of MA and non-verbal increases the correlation with achievement.

10.	Simple r, achievement with verbal	0.65
11.	Multiple r, achievement with (verbal + MA)	0.699
12.	Multiple r, achievement with (verbal + MA + non-verbal)	0.702

The properly weighted composite of the verbal group test and the MA yields a very high correlation with achievement. Non-verbal tests add very little to the combination. Whether the 0.05 increase over verbal produced by adding MA is worth while, is after all an

TABLE II.—CORRELATIONS FOR GRADE III[1]

	1 Dearborn 1	2 Dearborn 2	3 Dearborn 3	4 Dearborn total	5 Otis Primary	6 Haggerty	7 Meyers	8 Holley Vocabulary	9 National Intelligence Test A	10 National Intelligence Test B	11 National Intelligence Test Total	12 Illinois	13 Thorndike-McCall	14 Chronological Age	15 Stanford Mental Age	16 Reading Comprehension	17 Reading Rate	18 Arithmetic	19 Spelling	20 Composite Achievement	21 Approximate time (minutes)
Non-verbal tests																					
1. Dearborn 1	....	0.42	0.48	0.79	0.53	0.23	0.36	0.18	0.26	0.10	0.19	0.47	0.15	0.20	0.14	0.18	0.16	0.10	0.08	0.21	40
2. Dearborn 2	0.42	....	0.23	0.60	0.28	0.31	0.22	0.11	0.60	0.18	0.38	0.31	0.26	0.14	0.10	0.21	0.14	0.21	0.18	0.30	40
3. Dearborn 3	0.48	0.23	....	0.78	0.34	0.44	0.37	0.30	0.20	0.14	0.22	0.47	0.03	0.13	0.11	0.10	0.15	0.22	0.12	0.20	40
4. Dearborn total	0.79	0.60	0.78	....	0.52	0.59	0.45	0.26	0.41	0.19	0.33	0.48	0.24	0.20	0.15	0.20	0.18	0.18	0.13	0.30	120
5. Otis Primary	0.53	0.28	0.34	0.52	....	0.42	0.49	0.10	0.24	0.08	0.28	0.17	0.13	0.50	0.18	0.03	0.02	0.27	0.11	0.18	22
6. Haggerty	0.23	0.31	0.44	0.59	0.42	....	0.38	0.25	0.20	0.08	0.26	0.58	0.10	0.16	0.24	0.09	0.07	0.12	0.13	0.18	12
7. Meyers	0.36	0.22	0.37	0.45	0.49	0.38	.....	0.27	0.19	−0.04	0.15	0.36	0.02	0.15	0.20	0.02	0.07	0.08	0.16	0.19	10
8. Mean	0.40	0.29	0.37	0.52	0.43	0.40	0.38	0.21	0.30	0.10	0.26	0.41	0.13	0.21	0.16	0.12	0.11	0.17	0.13	0.22	164[2]
9. SD	0.10	0.07	0.09	0.07	0.09	0.11	0.08	0.07	0.14	0.07	0.07	0.13	0.09	0.12	0.05	0.07	0.06	0.06	0.03	0.05	
Verbal tests																					
10. Holley Vocabulary	0.18	0.11	0.30	0.26	0.10	0.25	0.27	....	0.65	0.59	0.63	0.76	0.58	−0.26	0.33	0.69	0.64	0.23	0.58	0.64	15
11. National Intelligence Test A	0.26	0.60	0.20	0.41	0.24	0.20	0.19	0.65	....	0.63	0.79	0.63	0.57	0.08	0.44	0.77	0.65	0.47	0.63	0.68	17
12. National Intelligence Test B	0.10	0.18	0.14	0.19	0.08	0.08	−0.04	0.59	0.63	.....	0.74	0.43	0.76	−0.33	0.38	0.67	0.70	0.38	0.64	0.60	16
13. National Intelligence Test A + B	0.19	0.38	0.22	0.33	0.28	0.26	0.15	0.63	0.79	0.74	....	0.61	0.80	−0.16	0.53	0.69	0.77	0.46	0.71	0.78	33
14. Illinois	0.47	0.31	0.47	0.48	0.17	0.58	0.36	0.76	0.63	0.43	0.61	....	0.38	−0.01	0.23	0.59	0.46	0.20	0.34	0.52	17
15. Thorndike-McCall	0.15	0.26	0.03	0.24	0.13	0.10	0.02	0.58	0.57	0.76	0.80	0.38	....	−0.16	0.16	0.88	0.83	0.38	0.45	0.70	30
16. Mean	0.23	0.31	0.23	0.32	0.17	0.25	0.16	0.64	0.62	0.60	0.68	0.56	0.62	−0.15	0.35	0.72	0.68	0.35	0.56	0.65	95[2]
17. SD	0.12	0.16	0.14	0.10	0.07	0.16	0.14	0.06	0.03	0.11	0.09	0.14	0.11	0.14	0.12	0.09	0.12	0.10	0.13	0.08	
18. MA	0.14	0.10	0.11	0.15	0.18	0.24	0.20	0.33	0.44	0.38	0.53	0.23	0.16	0.10	....	0.29	0.20	0.42	0.32	0.47	

[1] In computing the mean intercorrelations, the coefficients of part with whole and whole with part (National Dearborn) have been omitted.

[2] Totals, in which times for composite tests (National Intelligence Test and Dearborn) are included but once.

administration problem depending on the circumstances and needs of the school.

3. *The Relation of the Validity of the Test to Its Length.*—Fairly high correlations were found between the length of the non-verbal tests and the magnitude of the *r's* with the criterion in the case of the primary grades. In grade III, the correlations with composite of achievement will be used as a criterion for the reason that it is probably more reliable than any other. The correlations (Product Moment formula) of the length of the test (column 21, Table II) and the *r's* with achievement (column 20) are for non-verbal tests 0.76. The correlation between length of the verbal tests and the magnitude of the *r's* with achievement is 0.81. For both types of material, even with other things (methods of weighting in scoring, differences in validity of constituent exercises, etc.) varying as they do, the longer the test the higher the predictive value.

4. *Correlations with Achievement in Particular School Subjects.*—Columns 16 to 20 give the data, which are here summarized:

	Correlation of Mean Non-verbal	Correlation of Mean Verbal
1. With Reading Comprehension...	0.12	0.72
2. With Reading Rate.....	0.11	0.68
3. With Spelling.....	0.13	0.56
4. With Arithmetic.....	0.17	0.30
5. With Composite of Achievement.	0.22	0.65

Since the measures of achievement are largely verbal the fact that the verbal intelligence tests yield much higher correlations with them need not be surprising. Arithmetic is commonly judged to be less verbal than reading or spelling and the verbal intelligence tests, in fact, yield the lowest correlations with arithmetic, but the correlation of non-verbal with arithmetic is still lower. None of our material gives a useful prediction of this function. Reading and spelling, being decidedly verbal, yield high correlations with verbal group tests.

5. *Correlations with Chronological Age.*—Column 14, Table II, yields the following summary:

Mean *r*. non-verbal with age +0.21
Mean *r*. verbal with age −0.14

The data are in substantial agreement with the facts found in the case of the primary grades. Why non-verbal tests tend, much more than verbal, to be positively associated with age in the case of school grades, cannot be objectively determined by our data.

(*To be concluded in April*)

INTELLIGENCE IRREGULARITY AS MEASURED BY SCATTERING IN THE BINET SCALE[1]

J. E. W. WALLIN

Bureau of Special Education

(Subnormal and Delinquent Children)

Miami University

I shall attempt here to give a brief resumé of a number of studies, of unevenness in intelligence development as measured by scattering in the Binet-Simon scale which I have made during the last decade. As here employed, *scattering* signifies the number of tests passed in the Binet-Simon scale above the basal age. Our analyses of scattering thus far published have been confined to the 1908 and 1911 Binet-Simon scale, and have been based on 333 epileptics in an institution,[2] on 34 psychotics in institutions,[3] and on 1181 consecutive subjects of various types coming to a university and a public school psycho-educational clinic, classified both according to intelligence age and according to diagnostic category (determined to a considerable extent by grade of intelligence), 840 of whom were thoroughly tested and 341 less thoroughly tested.[4] All except two of the insane were adults. Thirty per cent of the epileptics were under 21 years of age, the youngest being 5 years old.[5] In the group of 840 school cases the average chronological age of the boys was 11.10, of the girls 10.87, and of both sexes 11.07; while the average intelligence age by the 1908 scale was 8.59 for the boys, 8.10 for the girls, and 8.45 for both; and by the 1911 scale 8.16 for the boys, 7.43 for the girls, and 7.95 for both. For the

---

[1] Presented, in extract, before the Section of Clinical Psychology of the American Psychological Association, December 28, 1921.

[2] "Experimental Studies of Mental Defectives," 1912, p. 22—an incomplete analysis which cannot be completed because of the inaccessibility of the original records.

[3] "Problems of Subnormality," 160f (of 1921 reprint).

[4] The Phenomenon of Scattering in the Binet-Simon Scale. *Psychological Clinic*, 1917, pp. 179–195.

Wide Range Versus Narrow Range Binet-Simon Testing. *Journal of Delinquency*, 1917, pp. 315–330.

A Further Comparison of Scattering and of the Mental Rating by the 1908 and 1911 Binet-Simon Scales. *Journal of Delinquency*, 1918, pp. 12–27.

An Analysis of Binet-Simon Records. *School and Society*, March 30, 1918.

[5] The distribution of the chronological ages for the children is given in "Experimental Studies of Mental Defectives," 1912, p. 102.

group of 341 the average chronological age was 10.4 for the boys, 10.1 for the girls, and 10.3 for both sexes.

Since the publication of the above studies we have finished an analysis of scattering by the same method[1] among 1025 consecutive

TABLE I.—DISTRIBUTION OF CHRONOLOGICAL AND INTELLIGENCE AGES

Ages	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	28	Total
Chronological:																				
Number boys	..	1	..	2	4	34	96	158	132	109	84	68	27	8	6	1	1	2	1	734
Number girls	..	1	..	...	3	16	38	41	44	47	40	34	9	10	3	3	..	2	..	291
Both	..	2	..	2	7	50	134	199	176	156	124	102	36	18	9	4	1	4	1	1025
Intelligence:																				
Number boys	5	19	32	65	156	162	153	85	32	17	3	1	0	1	..	..	..	..	..	731
Number girls	5	13	20	39	65	69	56	11	4	2	0	0	0	0	..	..	..	..	..	284
Both	10	32	52	104	221	231	209	96	36	19	3	1	0	1	..	..	..	..	..	1015[1]

Average chronological ages: boys 10.47; girls 10.96; both 10.61. The average intelligence ages are: boys 7.01; girls 6.71; both 6.92.

[1] A few cases have been omitted in this series because an exact intelligence age could not be assigned or because the records were inadvertently not transferred.

cases examined by means of the Stanford-Binet in a school psycho-educational clinic, when classified both according to intelligence age and diagnostic category. The distribution of the chronological and intelligence (Binet-Simon) ages for this group is shown in Table I.

The average chronological ages in the three groups of public school clinic cases do not differ significantly. But the average intelligence age is perceptibly lower in the Stanford, amounting to one year as compared with the 1911 scale, and about a year and a half as compared with the 1908 scale. The lower rating in the Stanford is probably due to the fact that this scale is more difficult, as shown by earlier analyses.[2]

THE RELATION OF SCATTERING TO GRADE OF INTELLIGENCE

Binet and Simon[3] asserted that the "defective child" is "inferior, not in degree, but in kind. The retardation of his development has not been uniform. . . . So far as certain faculties are concerned, he remains on the level of a younger child; but in respect to others, he is on a level with normal children of his own age. An unequal and imperfect development is consequently his specific characteristic.

[1] The detailed data will be published elsewhere.

[2] *E.g.*, "The Results of Retests by the Binet Scale." *Journal of Educational Psychology*, 1921, p. 399.

[3] "Mentally Defective Children," 1914, p. 13.

These inequalities of development . . . always produce a want of equilibrium, and this want is the differentiating attribute of the defective child." Doll[1] has raised this want of equilibrium to the importance of a pathognomonic sign of "potential feeble-mindedness." The component mental processes which determine the intellectual capacity develop uniformly in the former (the normal child) and not so uniformly in mental defectives. In the Binet tests the typical normal child has a basal year not more than one year below his chronological age and passes but a few tests beyond his chronological age. The potential feeble-minded, on the other hand, has the basal year more than one year below the chronological age or at least seriously below the total mental age rating and may have more than one basal year, that is to say, he 'scatters,' failing in tests one would expect a normal child of that age to pass, and succeeding in others not expected.

We have made a five-fold analysis of our data on scattering in relation to grade of intelligence, based on 2206 Binet records, and find little, if anything, in support of the above conclusions.

1. The greatest amount of scattering (based on the average amount of advance credit earned) is found among the normals in the Stanford and 1911 scales, and among the deferred in the 1908 scale, while the grade showing the least scattering is the imbeciles in the Stanford scale, the morons in the 1911 scale, and the normal in the 1908 scale.

2. If we confine the comparison to the extreme cases of scattering, we find that the proportion of subjects who earn 11 or more credits above the basal age (Table II), is greatest for the normal in the 1911 and Stanford scales, and for the deferred in the 1908, and least for the borderline in the 1911 scale, and for the imbeciles in the Stanford and 1908 scales.

3. Compared with the average amount of scattering found for all the subjects tested by each scale, it is found that in the Stanford scale the imbeciles and potential feeble-minded scatter less than the average (by 0.24 and 0.13 year, respectively), the morons slightly more (by 0.11 year), and the normals appreciably more (by 0.49) year. In the 1911 scale those who scatter less than the average are the imbeciles (by 0.13 year), the morons (by 0.13 year), and the borderline (by 0.03 year), while the normals scatter more (by 0.17 year). In the 1908 scale those who scatter less than the average are the imbeciles

[1] Preliminary Note on the Diagnosis of Potential Feeble-mindedness. *Training School Bulletin*, May, 1916.

TABLE II.—PERCENTAGE OF SUBJECTS PASSING ELEVEN OR MORE TESTS ABOVE THE BASE

	1908 scale			1911 scale			Stanford-Binet		
	Boys	Girls	Both	Boys	Girls	Both	Boys	Girls	Both
Normal	10.8	0.0	8.5	24.3	10.0	21.2	56.2	50.0	55.5
Retarded	7.5	14.2	8.9	15.1	14.2	14.9	39.1	0.0	31.0
Backward	3.6	14.8	6.2	18.2	8.1	15.6	43.4	34.7	42.3
Deferred	14.2	15.7	14.5	14.3	5.2	11.1	40.0	16.6	33.3
Borderline	3.7	12.5	5.7	6.1	4.1	5.7	44.3	41.1	43.6
Potential feeble-minded	....	....	....	....	....	....	36.6	15.3	30.9
Borderline and potential feeble-minded	....	....	....	....	....	....	42.5	34.0	40.4
Morons	9.6	21.5	13.8	6.4	9.8	7.6	43.0	44.0	43.4
Potential morons	....	....	....	....	....	....	46.9	44.9	46.0
Morons and potential morons[1]	....	....	....	....	....	....	44.7	44.4	44.6
Imbeciles	0.0	2.2	0.9	10.1	13.6	11.6	30.1	20.7	26.2
All	5.6	13.1	7.7	13.7	9.3	12.5	41.6	34.7	39.7

[1] In the 1908 and 1911 classifications the potential morons are included among the morons, and the potential feeble-minded among the borderline.

(by 0.02 year), the borderline (by 0.06 year), and the normals (by 0.25 year), while the morons scatter more (by 0.14 year).[2]

4. If we group together as *normal* all of those diagnosed as above borderline intelligence (*i.e.*, the normal, retarded and backward), as *subnormal* all those diagnosed as borderline and lower, and as *feeble-minded* the morons (including potential morons) and imbeciles, we find that in the 1911 scale the scattering is the least for the feeble-minded, and in the 1908 scale for the normals, while it is greatest for the normal group in the 1911 scale. The difference between the feeble-minded and normal group amounts to 0.16 year in both scales. The averages are the same for the subnormal and feeble-minded group in the 1908 scale, and for all groups in the Stanford (the difference amounting to only 0.01 year).

5. We may, finally, consider the number of ages from and including the base in which one or more tests were passed. The morons and potential morons pass tests in the greatest number of ages in the Stanford scale, the backward in the 1911, and the deferred in the 1908, while the retarded pass tests in the smallest number of ages in the

[2] The average amount of scattering for all the subjects is 1.02 years in the 1908 scale, 1.39 years in the 1911 and 1.63 years in the Stanford.

Stanford, (the "retarded" grading slightly under normal with an average IQ of 91), the imbeciles in the 1911, and the normals in the 1908.

It is evident from the foregoing that the conclusions would differ somewhat according to the scale followed. We believe, however, that some of the significant discrepancies may be cleared away by a study of, first, the average intelligence age level of the subjects in each diagnostic category in each scale and, second, by an analysis of the extensiveness of the testing of the subjects in each category. It is evident that groups of subjects with a high intelligence level could not scatter very much in the 1908 scale because of the small number of tests in ages 10 to 13, and the complete absence of tests above age 13. Moreover, in a previous comparison of scattering (in the 1908 scale) among subjects extensively tested and those sketchily tested, we found that the average number of advanced credits earned was invariably larger for the extensively tested group in every Binet-Simon age admitting of comparison. The pupils given a wide-range test earned from 0.75 of a year to 1.75 years more extra credits than those given a narrow-range test. Curiously the ratio of the number of tests passed to the number given was greater for the extensively tested than for the sketchily tested showing that a disproportionate amount of the credits came from the higher ages. On the basis of the experimental findings we were led to conclude that "the amount of credit

TABLE III.—AVERAGE INTELLIGENCE AGE (BINET-SIMON) OF SUBJECTS IN EACH CLASSIFICATION

	Stanford	1908 scale	1911 scale
Normal	9.34	9.26	8.77
Retarded	9.31	8.95	8.47
Backward	8.27	9.25	8.85
Deferred	5.93	7.07	6.62
Borderline	7.04	8.96	8.49
Potential feeble-minded	6.66		
Borderline and potential feeble-minded	6.95		
Morons	7.27	8.32	7.85
Potential morons	6.00		
Morons and potential morons	6.71		
Imbeciles	4.51	5.69	4.74
Average	6.92	8.45	7.95

earned depends upon the extent of the testing and not upon the grade of intelligence of the pupils."[1]

An analysis of the figures in Table III suggests the scattering among the normal pupils in the 1908 scale was small because of their high intelligence level (the same explanation would apply to the backward in this scale whose intelligence level was the same and who scattered less than the average). Forty-two per cent of the normals had a base of from X to XIII.

On the other hand, the extensiveness of the testing (see Table IV), helps to explain why the moron average of scattering is somewhat above the general average in the 1908 scale, and why scattering is greatest among the deferred in this scale, who also pass tests in the largest number of ages from and including the base. It helps to explain why the moron average is slightly above the general average in the Stanford scale, why the morons and potential morons in this scale pass tests in the largest number of ages from and above the base, and why the retarded (and backward) pass tests in the smallest number of ages.

TABLE IV.—AVERAGE NUMBER OF TESTS GIVEN ABOVE THE BASAL AGE

	In 1908 scale	In Stanford-Binet
Normal	8.7	19.5
Retarded	11.6	14.0
Backward	12.1	19.8
Deferred	18.4	22.9
Borderline	13.4	22.7
Potential feeble-minded	....	23.8
Borderline and potential feeble-minded	....	23.0
Morons	15.6	24.8
Potential morons	....	28.2
Morons and potential morons	....	26.3
Imbeciles	18.6	23.8
All	13.8	22.8

From the above analyses we only seem justified in concluding that normal pupils scatter most and imbeciles least. Certainly there is no warrant for the assumption that "unequal development," "lack of uniformity," or "scattering" in intelligence is the "specific charac-

[1] *Journal of Delinquency*, 1917, 315f.

teristic," the "differentiating attribute," the pathognomonic sign of feeble-mindedness or potential feeble-mindedness. In fact, it seems to be the very want of irregularity, at least so far as general intelligence is concerned as determined by scattering in the Binet scale, which characterizes the lower grades of the feeble-minded. The conclusion here reached is in harmony with our earlier findings and with the findings of Pressey, who has studied irregularity among various types of subjects in the Binet scale, by different methods,[1] and Mathews, who has studied the problem in a group of delinquents.[2]

The Relation of Scattering to Intelligence Age Level

The smallest amount of scattering occurs in ages 2, 3, and 4, and in ages 10 to 12 in the 1908 scale. The data from the latter ages are probably unreliable because of the limits of the scale at the upper end. In the 1911 scale the corresponding ages are 2, 12, and 13 (the scattering being, again, artificially limited in the latter two ages because of the limited number of tests), and in the Stanford scale ages 3 and 10.[3] The greatest amount of scattering comes in ages 6 and 7 in the 1908 scale, in ages 11, 6, 9 and 10 in the 1911 scale, and in ages 11, 8 and 5 and 6 in the Stanford (ignoring ages above 11 because of the fewness of the subjects). The results are discrepant in the different scales, indicating, again, the danger of attempting to draw positive conclusions from one scale of tests. If we are justified in drawing any conclusions at all, it would be that the smallest amount of scattering occurs in the lowest mental age levels, and possibly that the largest amount tends to occur in the middle range of ages. This conclusion does not entirely harmonize with other findings. Pressey's tabulation shows that the scattering is about the same in the IQ range from

[1] Pressey, S. L.: Irregularity on a Mental Examination as a Measure of Its Reliability. *Psychological Clinic*, 1919, 236f.

Irregularity on a Psychological Examination as a Measure of Mental Deterioration. *Journal of Abnormal Psychology*, December, 1918.

The Distinctive Features in Psychological Test Measurements. *Journal of Abnormal Psychology*, 1917, 3f.

A Comparison of a Girls' Reform School, Attendants at a State Hospital for the Insane and Public School Children, by Means of Certain Tests of Intelligence. *Journal of Criminal Law and Criminology*, 1921, 258f.

[2] Mathews, Julia: Irregularity in Intelligence Tests of Delinquents. *Journal of Delinquency*, 1921, 355f.

[3] It should be pointed out that the differences referred to in this and other sections are sometimes almost negligibly small.

—76 to 125, and perceptibly higher in the range above 125.[1] On the other hand, in Mathew's data for delinquents the scattering is least for IQ's below 76 among the boys, and for IQ's below 76 and from 110 to 125 for girls.[2] Whatever significance, if any, differences of scattering may have in relation to intelligence age level, it is well to point out that the difference is quite marked between the age levels showing the least and the greatest amount of scattering, amounting to a year in the 1908 scale, a year and a third in the Stanford, and almost a year and two-thirds in the 1911. These differences are considerably greater than the corresponding differences between the intelligence categories.

The Relation of Scattering to Neurotic, Psychopathic and Delinquent Types

We shall compare (1) the average number of advance credits earned by each of these groups and by all of them combined into one group[3] with the corresponding average for all the subjects tested by each scale, for the combined group of normals ("normal," "retarded" and "backward") and for the limited group of normals, *i.e.*, those strictly diagnosed as normal; and (2) the percentage of subjects in each of these unstable groups and among all those tested and among those strictly diagnosed as normal who passed more than 10 advance tests in each scale.

Neurotics

1. The average amount of scattering for the neurotics was 0.01 year less than for all the subjects in the 1911 scale, and 0.24 year less than for all the subjects in the 1908 scale, but 0.17 year greater than for all the subjects in the Stanford scale. Compared with the combined normal group, the scattering was 0.16 year less for the neurotics in the 1908 scale, 0.06 year less in the 1911, and 0.19 year more in the Stanford. Compared with the limited normal group, the neurotics scattered 0.02 year more in the 1908 scale, 0.18 year less in the 1911 scale, and 0.31 year less in the Stanford.

2. Based on the percentage of subjects who passed 10 or more

[1] *Psychological Clinic*, 1919, p. 236.

[2] As before.

[3] Number of neurotics, 155; of psychopaths, 22; of delinquents, 352; total, 529. The diagnosis of these types was based on the physicians' reports, the psychological finding, and the personal history (social and school records). The delinquencies include such offenses as truancy, disorderliness, lying, stealing, and bad sex practices.

advance tests the ratio was 7.7 per cent less among the neurotics in the 1908 scale, but 9 per cent greater in the 1911 scale and 8.6 per cent greater in the Stanford. Compared with the limited normal group, the proportion for the neurotics was 8 per cent less in the 1908 scale, the same in the 1911 scale and 7.2 per cent less in the Stanford.

Psychopaths

1. The median amount of scattering among the psychopaths exceeds the general average by 0.10 year in the 1908 scale, by 0.03 year in the 1911 scale, and by 0.27 year in the Stanford. Compared with the combined normal group, the psychopathic excess amounts to 0.19 year in the 1908 scale, 0.04 year in the 1911, and 0.29 year in the Stanford. Compared with the limited normal group, there is a psychopathic excess in the 1908 scale amounting to 0.37 year, and a psychopathic deficiency in the 1911 and Stanford scales amounting to 0.14 year and 0.21 year, respectively.

2. Compared with all the subjects, the percentage who passed over 10 advance tests is greater for the psychopaths by 6.5 per cent in the 1908 scale, by 30.6 per cent in the 1911, and by 13.6 per cent in the Stanford scale. Compared with the limited normal group, the proportion among the psychopaths is 6.2 per cent and 21.8 per cent higher in the 1908 and 1911 scales, respectively, and 1.6 per cent less in the Stanford.

Delinquents

1. The average amount of scattering among the delinquents is 0.17 year less than among all the subjects in the 1908 scale, but exceeds the general averages by 0.04 year in the 1911 scale, and 0.03 year in the Stanford scale. Compared with the combined normal group, the delinquents scatter 0.09 year and 0.01 year less in the 1908 and 1911 scales, and 0.05 year more in the Stanford. Compared with the limited normal group, the delinquent average is 0.09 year greater in the 1908 scale, 0.13 year less in the 1911, and 0.45 year less in the Stanford.

2. Compared with the whole group, the proportion of delinquents who passed over 10 advance tests was 3.8 per cent less in the 1908 scale, but 5.1 per cent greater in the 1911 scale and 15.2 per cent greater in the Stanford. Compared with the limited normal group, the proportion among the delinquents was uniformly less in all the scales, amounting to 4.1 per cent in the 1908 scale, 3.7 per cent and 0.6 per cent in the Stanford.

Unstables

Many writers find it impossible to draw any clearly definable, unambiguous distinction between neurotics and psychopaths, while others are inclined to consider many, if not most, delinquents as psychopaths or neurotics. It will be of interest here to average the results for all these subjects and treat them as a single group of "unstables." When so treated we find: (1) the scattering is 0.17 year less among the unstables than among all the subjects in the 1908 scale, but 0.02 year more in the 1911 scale and 0.09 year more in the Stanford scale. Compared with the combined normal group, the scattering among the unstables is 0.09 year less in the 1908 scale, 0.02 year less in the 1911 scale, and 0.11 year more in the Stanford. Compared with the limited normal group, the unstable average is 0.09 year greater in the 1908 scale, but 0.14 year less in the 1911, and 0.39 year less in the Stanford. (2) The proportion of unstables who pass over 10 advance tests is 4.1 per cent less in the 1908 scale, but 6.5 per cent more in the 1911, and 12.5 per cent more in the Stanford, when compared with the figures for all the subjects. Compared with the limited normal group, the proportion is less for the unstables in all the scales, amounting to 4.4 per cent in the 1908 scale, 2.3 per cent in the 1911 and 3.3 per cent in the Stanford.

Here, again, the results differ according to the particular scale and according to the particular criterion employed. The average for the unstables exceeds that for the feeble-minded group; it possibly tends to exceed the general average slightly, although the tendency is by no means uniform; while among the unstable group the scattering is slightly greater for the psychopathic in nearly all the comparisons. But it is a question whether this peculiarity can be elevated to the rank of a pathognomonic sign of psychopathy, as has been done by Mateer,[1] who has accepted "more than four years of scattering above the basal year as an indication of psychopathy." We have found some individuals among all types—normal, subnormal, feeble-minded, delinquent, neurotic, psychopathic—who scatter very greatly, and others who scatter very little. Mathews records the same observation, "While many individuals who show great scattering in their tests are recognized as unstable, neurotic or psychopathic in their make-up, there are others in whom these tendencies are quite as pronounced who

[1] Mateer, Florence: The Future of Clinical Psychology. *Journal of Delinquency*, 1921, p. 283f.

give very even tests." Some recent writers have emphasized the importance of accompanying the mental age rating with a statement of the amount of scattering. That may be advisable, but the significance of little or much scattering has not yet been made clear by the studies which have thus far appeared. It is not yet certain whether scattering can be used as a pathognomonic sign of any type of mental defect, although our previous analyses seem to have shown that epileptic and psychotics as groups scatter more than any other groups—a conclusion apparently also reached more recently by Pressey. Nevertheless, we have examined epileptics and psychotics who scatter very little. Irregularity in intelligence development is not an indispensable condition or accompaniment of the epilepsies or the psychoses, although it is a frequent complication, possibly chiefly in the demented stages.

The Relation of Scattering to Sex

In the 1908 scale, the girls scatter more than the boys in all the diagnostic classifications except two, the difference between the averages for all the boys and girls amounting to 0.21 year. If the comparison is confined to those who do over 10 advance tests, the proportion is greater for the girls in all the classifications except one, while the proportion for the whole group who thus scatter is 7.5 per cent greater among the girls than among the boys. On the other hand, the boys earn more advance credits in every diagnostic classification in the 1911 and Stanford scales (using the complete figures), the boys' average amount of scattering exceeding that of the girls by 0.10 year in both scales. If we confine the comparison to the subjects passing more than 10 advance tests, the proportion is larger for the boys in all except two classifications in the 1911 scale, and in all except one in the Stanford. On the average, the proportion of boys who thus scatter exceeds the corresponding proportion of girls by 4.4 per cent in the 1911 scale and by 6.9 per cent in the Stanford. Thus the girls consistently vary more in the 1908 scale, and the boys in the other two, which are usually recognized as being more accurate than the 1908. This might be considered as slight confirmation of the view that the male sex is more variable than the female. Yeung,[1] in a recent examination of a small number of Chinese by the Stanford-Binet, found that the girls were only 0.63 as variable as the boys according to the Pearson coefficient of variability.

[1] Yeung, Kwok Tsuen: The Intelligence of Chinese Children in San Francisco and Vicinity. *Journal of Applied Psychology*, 1921, p. 267f.

Difference in Scattering in the Different Scales

Ordinarily we should suppose that the scale which shows the greatest amount of scattering is the most inaccurate. We should expect it to contain a larger number of improperly placed tests. It is, however, difficult to compare the scattering in the scales we have used because of the lack of tests in the upper ages of the 1911 scale, and especially the 1908 scale, because of the variation in the number of tests in the different ages in the 1908 scale, the difference in the method of obtaining the base in this scale, because of the larger number of tests, and the absence of tests in certain ages in the upper part of the Stanford scale, because of the unequal value of the tests in the Stanford scale above age 10, and because the extent of the testing may not have been equal in the different scales. Bearing these difficulties in mind, we find that the scattering is considerably greater in the Stanford scale, the average excess amounting to 0.24 year as compared with the 1911 scale, and 0.61 year as compared with the 1908. We have been aware of this peculiarity or weakness of the Stanford scale from our earliest experience with it. We have been struck by the unusually low bases which subjects frequently obtain in this scale because of failures on single tests in various ages which seem to be of more than average difficulty. It is frequently necessary to test all the way down to ages four or five in the case of subjects grading eight years or higher in intelligence, in consequence of which it is necessary to give a large number of tests in this scale in order to be fair to the subject. As a matter of fact, the testing was considerably more extensive in the Stanford than in the 1908 scale, for which we have comparable data, as indicated by the following figures, which show the number of tests given above the base in each category in each scale.

Table V.—Number of Tests Given above the Base in each Diagnostic Category

	Normal	Retarded	Backward	Deferred	Borderline	Morons	Imbeciles
1908	8.7	11.6	12.1	18.4	13.4	15.6	18.6
Stanford	19.5	14.0	19.8	22.9	23.6	24.8	23.8

A STUDY OF HIGH SCHOOL SPELLING MATERIAL II

JOHN A. LESTER

The Hill School, Pottstown, Pa.

(*Continued from February Journal*)

4. *Groupings for Purposes of Teaching.*—The reader will observe that the purpose of the investigation was not in the main to seek causes but to classify phenomena for the purpose of economical teaching. The aim was not to explain *why* these words are misspelled, but to observe *how* they are misspelled, and to group together the misspellings which exhibit a kindred nature. In certain cases, as for instance the fourth class in Table IV below, this classification indicates the cause at the root of the misspellings indicated. But, in general, the aim is to show the nature, degree, and percentage of certain forms of misspelling. The results of this classification are given in the form of a table. Each column for convenience is numbered and corresponding explanatory notes are added.

1. *Word-compounding.*—Under mistakes of word-compounding are classified all forms whether written solid, separate, or hyphenated (*twentyfive, him self, with-out*), as have no support in good usage, and such as plainly alter the meaning intended by the writer (*e.g., a white bearded barber*). The prevalence of this kind of inaccuracy in writing is revealed only from accumulated records of errors made in free composition. But few of the mistakes in compound adjectives preceding their nouns (*e.g., high powered engine*), though occurring with great frequency as a class, will occupy a high place in the arrangement of individual frequencies. They are individual, and occur in response to a special need. To determine the relative amount of this kind of misspelling, the sum total of the recorded mistakes in word-compounding was computed, compared with the sum total of recorded misspellings of all kinds, and found to constitute 15.9 per cent of the errors in spelling in free composition. This figure, 15.9, appearing in the table in parentheses, indicates the weight of this error more exactly than the percentage 6.2, which indicates the mistakes of this nature occurring in the most commonly misspelled 775 words.

A detailed examination of the mistakes of word-compounding shows that the tendency to separate is 15 times stronger than the tendency to combine in the writing of the students whose work is considered in this investigation.

TABLE IV.—CLASSIFICATION OF THE 10,853 TYPE MISSPELLINGS OF THE 775 WORDS MOST FREQUENTLY MISSPELLED

	1	2	3	4	5	6	7	8	9	10
Nature of the error	Word-compounding	Prefixes and suffixes	Confusion of words similar in sound or appearance	Mispronunciation	Apostrophe in possessives	Analogy	Doubling or singling of consonants	Final single silent *e* before a suffix	Mistakes which would be avoided by knowledge of some common Latin roots	Substitution of a vowel or vowel combination for one of similar sound
Percentages of the total misspellings	(15.9) 6.2	15.4	13.3	12.1	(8.2) 2.9	7.6	6.4	3.7	3.6	3.6

TABLE IV.—*Continued*

	11	12	13	14	15	16	17	18	19	20	21
Nature of the error	Final single consonant before a suffix	Substitution of a consonant or a consonant combination for one of similar sound	Internal modification: attraction, absorption, transposition	ei, ie	Capitalization	Omission or addition of a final silent *e*	Approximation to phonetic spelling	Apostrophe used for contraction	Final *y* before a suffix	Spellings in French form	Doubtful
Percentages of the total misspellings	3.4	3.4	3.4	3.2	(2.1) 0.4	2.0	1.4	1.3	1.3	0.9	4.5

2. *Mistakes in Prefixes and Suffixes.*—In this classification, misspellings in the *form* of prefixes and suffixes, denoting ignorance of what they are, were differentiated from misspellings made in the junction of prefixes and suffixes with the stem, indicating often ignorance of the general principles of word formation. Mistakes in the two noun suffixes *-ance* (*-ence*), *-er* (*-ar*, *-or*), and in the suffix *-ent* (*-ant*) account for more than half of all errors of this kind; suffixes are in general misspelled more often than prefixes; and the aggregate of mistakes in the form of suffixes and in the composition of prefixes amounts to 84 per cent of all the mistakes of this nature.

3. *Confusion of Similar Words.*—Under this head were classified mistakes not only in homonyms (*hear*, *here*), but in words which show such similarity of form as to lead to occasional substitution (*formally*, *formerly*). The great amount of misspelling attributable to the confusion of similar words is due to the high frequency of common words like *too*, *principal*, *there*, *led*, whose forms have not been firmly fixed in the minds of children in the elementary grades.

4. *Mispronunciation.*—There is a danger of classifying under this head any word whose misspelling appears to indicate a pronunciation other than the normal, but it is quite practicable to test the evidence of the records in the matter of pronunciation. Words whose common misspellings seem to indicate a common mispronunciation may be introduced into a typewritten piece of connected prose, and the pronunciation of the words observed when, without preparation, the prose is read aloud. This test proves for instance that while in 29 recorded misspellings of *thought*, 27 take the form of *though*, and all the recorded misspellings of *officer* take the form *officier*, neither *officier* nor *though* indicates a mispronunciation. On the other hand, it proves that the misspelling *atheletic* which occurs 32 times out of the recorded 35 misspellings of this word, is a true record of a common misspelling. This test was applied to all the words whose misspellings were in whole or part assigned to mispronunciation, and no word was included without evidence of (1), examples, few or many as the case was, of written forms indicating a mispronunciation; and (2), actual observation of such mispronunciation among students of seventeen to eighteen. The percentage indicates the amount of misspelling which is directly traceable to incorrect and slovenly habits of speech.

5. *Apostrophe in Possessives.*—Under this head is considered the omission, intrusion, or misplacement of the apostrophe denoting possession. It is clear, however, that no accurate measure of the

degree of this kind of misspelling in free composition can be obtained by taking its frequency in the 775 words most commonly misspelled. Hence all recorded misspellings of this class were computed, and the relative weight of this error (8.2) determined by comparing this total with the total of all misspellings of whatever kind.

But should such mistakes be regarded as misspellings rather than evidence of ignorance of grammar? Careful examination of the work of 560 candidates who wrote papers in 1918 and 1919 showed that 76 per cent of the books which contained errors in possessives also contained evidence elsewhere that the rules for the apostrophe to denote possession were known to the writers; and that the errors were therefore properly to be considered lapses, or errors of inattention.

Approximately one misspelling out of every twelve is a mistake in the form of the possessive.

6. *Analogy.*—This error differs from (3) above, in that here the form written is not a recognized word; the characteristic of the erroneous part of it bears an analogy to a recognized word or to part of one. In many cases this analogy produces reciprocal errors. For example, 85 per cent of the mistakes in *absence, sense,* take the forms *absense, sence;* 90 per cent of the mistakes of *always, all right* take the form *allways, alright;* 92 per cent of the mistakes in *choose, lose* take the forms *chose, loose.*

7. *Writing Single Consonants Double, Double Consonants Single.*—Under this head are classified mistakes of the nature indicated which do not fall into any other class.

8. *Final "e" Before a Suffix.*—Under this head are classified the mistakes in eliding or retaining a single silent final *e* before a suffix. The rule is commonly stated in three parts, as follows: A word ending in single silent *e* drops the *e* before a vowel. The silent *e* is retained before a consonant. The silent *e* is also retained to keep the soft sound of *c* and *g* before *a* and *o*. The percentage 3.7 indicates the applicability of this threefold rule.

9. *Common Latin Roots.*—Under this head are classified such misspellings as would obviously disappear in the light of a little accurate knowledge of Latin and of Latin roots. Arbitrary judgment was used in determining whether certain misspellings, such, for instance, as *comparitively,* should be placed under this head or another. These cases were always decided from the point of view of the practical teaching problem involved. The decreasing importance of Latin in the school curriculum was considered, and all cases such as the one men-

tioned (*comparitively*) which could be referred to both mispronunciation and to ignorance of Latin were referred to the former class. Under the head of ignorance of Latin might be placed many of the nouns and adjectives ending in *-ance, -ence, -ant, -ent.* But the same consideration of the meager knowledge of Latin which can be assumed to be practically available to the high school student in his spelling of English words, and the peculiar difficulties which characterize these endings, made it seem advisable to consider them separately. In other cases arbitrary judgment had to be used to decide whether a given misspelling (*e.g. innocent*) should be referred to mistakes in prefixes or to ignorance of Latin. Again where no knowledge of Latin can be assumed it is simpler to correct such an error by comparison with words similar in prefix, than with words similar in stem.

But when all such eliminations are made, there remain a number of misspellings which are most easily clarified by reference to the origin of the words in 42 common Latin roots. Examples are such forms as *benefit, ameteur, decend, descide, immagine, opperation, predjudice, volenteer.*

10. *Writing Another Vowel or Vowel Combination for One of Similar Sound.*—Under this head are classified mistakes of the nature indicated which do not fall into any other class. A number of misspellings of this general class are grouped with mispronunciations; others in which the prefixes (particularly *de-, re-*) or suffixes (particularly *-able, -ance, -al, -er, ious*) are written in incorrect form are grouped with mistakes of prefixes and suffixes; others are grouped with final *y*. Examples of the present group are such misspellings as *committy, privelege, endevor seveer, competetive, mineature, apperance.*

11. *Final Single Consonant before a Suffix.*—Under this head are classified mistakes of non-doubling or improper doubling of the consonant before the suffix which begins with a vowel. The common rule applicable in such cases may be expressed as follows. A monosyllable or a word accented on the last syllable, if it ends in one consonant preceded by one vowel, doubles the final consonant when a suffix beginning with a vowel is added. In other cases the consonant is not doubled. The rule applies with one exception in the list of 775 words.

12. *Substitution of a Consonant or a Consonant Combination for One of Similar Sound.*—From this classification is excluded as classified elsewhere mistakes of doubling or singling consonants; consonant mistakes in the junction of prefixes and suffixes with stems; several

mistakes of analogy; and several mistakes which would be removed by a knowledge of Latin roots.

Of the mistakes which were classified here, the writing of *sc* for *s* or *c* and the writing of *xs* for *x* (or vice versa) formed the largest group.

13. *Internal Modification: Attraction, Absorption, Transposition.*—Here were classed misspellings which show an expansion, contraction, or rearrangement of elements already existing in the proper form of the word. These internal changes may be classified as errors of attraction, in which an element already present either redundantly repeats itself, or else displaces another similar element (*accostom, athlectic, convienient*); errors of absorption, in which when two similar elements are present, one is assimilated to the other (*convient, irrestible, enthuiastic*); and errors of transposition, in which the elements suffer no expansion or contraction, but are merely rearranged (*entierly, Euorpean, gaurd, villian*). The amount of the misspellings due to each of these three processes is about the same.

14. *"ei," "ie."*—Here are classified all misspellings which would be rectified by the mechanical application of the rule which follows. When the diagraph *ei, ie* is sounded like *a* in *fate*, or when, having the sound of *e* in *be*, it follows *c*, write *ei*. In other cases write *ie*. There are three exceptions to this rule in the list of 775 words.

15. *Capitalization.*—Mistakes in capitalization are naturally widely scattered. To obtain the relative weight of this error, the sum of all recorded mistakes in capitalization was compared with the total number of misspellings in the entire mass of writing. The most usual form of the mistake is the omission of the capital in the writing of proper adjectives (*french, indian*).

16. *Omission of Addition of a Final Silent "e."*—Examples of omission are such forms as *glimps, Main, medicin, practis.* Examples of addition are such forms as *ascende, dependente, prefere, stomache.* Mistakes of omission are the more numerous in the proportion of 3 to 2.

17. *Approximations to Phonetic Spelling.*—Here are classified attempts, as it were, at phonetic spelling, usually characterized by the omission of a silent letter, or the substitution of one which more nearly represents the sound. Examples are *cubboard, curtesy, evrybody, gard, Wensday.*

18. *Apostrophe in Contractions.*—The mistake of misplacing the apostrophe in contractions is more frequent than the mistake of omitting it.

19. *Final "y."*—Under this head are classified mistakes in final *y*

when a suffix follows, or when in plurals and verbs it is changed to *-ies, -ied.* The rule as commonly expressed has two parts, as follows. Final *y* preceded by a consonant becomes *-ies* in noun plurals, and *-ies, -ied* in verbs. Final *y* preceded by a consonant becomes *i* before any suffix which does not begin with *i.* There are no exceptions to this rule in the list of 775 words.

20. *Misspellings in French Form.*—A small number of words are frequently misspelled by the substitution of the French form of the word or an approximation thereto. The evidence for this classification is (1) that the French form predominates in the misspellings of these words, and (2) that the misspellings are often observed to occur after the study of French has begun. Seventy-seven per cent of the misspellings of *affairs* take the form *affaires;* 91 per cent of the misspellings of *minute* take the form *minuit* or *minuite;* all the recorded misspellings of *cover, officer,* take the forms *couver, officier.*

Of the formal rules usually presented in the teaching of spelling only four appear in the above classification. They occur in columns 8, 11, 14, and 19. Conjointly, they apply to some 12 per cent of the observed misspellings. Other rules usually taught were examined to determine their value in the teaching of this material. These rules were taken from four spelling books, one of which prints 12 rules, another 14 and the third, 27. Since none of the rules examined had an application amounting to more than one-half of one per cent of the misspellings actually made, they are not included in the classification which is likely to prove of use in the teaching of this material.

A report of the results obtained from a system of presentation and drill based on the foregoing study lies beyond the scope of this paper. The use of such a system under classroom conditions shows beyond question that greatly increased efficiency in the process of learning to spell may be obtained with an accompanying decrease in the time and labor expended.

Conclusions

1. The prevalence of lapses shows that the teacher of spelling must apply himself (1) to the formation of habits of accuracy and attention in writing and in reviewing what has been written; and (2) to the imparting of a body of knowledge.

2. Since some 25 per cent of the misspellings of high school and preparatory school graduates are misspellings of derivatives, the teaching of material drawn up on a dictionary basis will not be adequate.

3. If the critical points of the words to be presented are definitely known, and if the material is presented with insistent emphasis upon these critical points, at which most misspellings originate, a great gain in time and efficiency may be effected.

4. The rules usually taught have a very small ratio of applicability to the material to be learned. The four commonest rules, relating respectively to final silent *e* before a suffix, final single consonant before a suffix beginning with a vowel, *ei* and *ie*, and finally *y* before a suffix, cover conjointly less than 12 per cent of the misspellings recorded.

5. The five largest classes of misspellings, together including nearly 65 per cent of the whole are, in order of importance, the following: (1) mistakes in word-compounding, (2) mistakes originating in prefixes and suffixes, (3) confusions of words similar in sound or in appearance, (4) mistakes traceable to mispronunciation, (5) mistakes in the use of the apostrophe to denote possession.

6. The most direct means of gaining economy and efficiency in the teaching of spelling would appear to be twofold. (1) To teach material which the students concerned *do not know;* in other words to select material not on the basis of word frequencies in adult vocabularies, but on the basis of word frequencies in the misspellings actually occurring in free written composition. (2) To present that material with insistent emphasis on those critical points in the words presented which, in the material studied above, cause nearly 77 per cent of the entire mass of misspellings.

Books and Articles Referred to

Ayres, Leonard P.: "A Measuring Scale for Ability in Spelling." New York City Division of Education, Russell Sage Foundation, 1915.

Bawden, H. Heath: A Study of Lapses. *Psychological Review*, Monograph Supplements, Vol. III, No. 4.

Cook, W. A. and O'Shea, M. V.: "The Child and His Spelling." The Bobbs-Merrill Company, Indianapolis, 1914. (In Childhood and Youth Series, ed. M. V. O'Shea.)

Gregory, B. C.: The Rationale of Spelling. *Elementary School Teacher*, 8, pp. 40–45.

Hollingsworth, Leta B. and Winford, C. Amelia.: The Psychology of Special Disability in Spelling. New York City, Teachers College, Columbia University, 1918. (Teachers College, *Columbia University Contributions to Education*, No. 88.)

A COMPARATIVE STUDY OF A BORDER LINE DEFECTIVE AND A NORMAL CHILD OF THE SAME MENTAL AGE

LAURA REMER

Stanford University

Introductory note by L. M. Terman.—During the last five years, several students of educational psychology at Stanford University have made minor studies of the educability of children testing at various degrees below normal. One of these studies, the results of which are not yet ready for publication, involved a year of tutoring, two to three hours daily, of an 18-year-old girl of about 70 IQ. The study was made by Mrs. Gertrude Bell, a professor in the San Diego State Teachers College. Mrs. Bell did the tutoring herself, using the most ingenious methods she could devise. The purposes of the study were: (1) to find out how much the subject's school achievement could be improved, as measured by elaborate educational tests taken before and after treatment; and (2) to find out whether such intensive educational treatment would affect the IQ. The net results of the experiment showed a large improvement in subject matter achievement, amounting on the average to nearly two grades advance for the year. On the other hand, the IQ showed no change except for the slight improvement which would be expected to result from several repetitions of the test. Even specific coaching on types of material analogous to (but not identical with) that composing the Stanford-Binet scale had only negligible effect on the IQ.

The present study, by Miss Laura Remer, is considerably less extensive than that of Mrs. Bell, but it is offered for publication because of the scarcity of carefully made and accurately reported observations on individual cases of low IQ. Although not all subnormal children of a given IQ show the same degree of inability to master school subject matter, the case described in this article gives a concrete and fairly accurate indication of what may be expected educationally, from the typical 8-year-old having 6-year intelligence. Incidently, it gives a rather vivid picture of the differences between 100 IQ and 70 or 75 IQ. We need similar comparisons, even more searching, of the differences between 100 IQ and 140 IQ or higher.

LEWIS M. TERMAN.

Grace M. was 8 years and 7 months old when she was brought to Stanford University for a mental examination because she was making no progress in learning to read. She was given the Stanford Revision of the Binet-Simon tests with results as follows: Mental age, 6 years, 6 months; IQ 75. This placed her in the border-line group between dullness and feeble-mindedness.

She had started to school when she was 6½ years old. She was now beginning her third year in school and although her teacher was giving her extra help she was not learning to read. She had never been placed in a special group for backward children.

Physically, she appears quite normal. She has never had a serious illness. In personal appearance she is attractive, bright-eyed and often vivacious.

The mother is a Hungarian and claims to have come from a family of very good standing. She is extremely temperamental and easily excited to the point of hysteria. The father, an epileptic, is an uneducated American. There are four children all of whom, except Grace, test between 90 and 96 IQ.

The members of the family took a very unsympathetic attitude toward Grace. They were unable to control her in any way. In the neighborhood she was considered a vicious child. At school she could not study and would not let anyone around her study in peace. Habits of lying, stealing, and cunning deceit were becoming rapidly fixed.

In deciding what action to take the following facts were considered. Most children begin to learn to read when they are mentally 6 years old or younger. Grace was mentally 6 years and 6 months. Accordingly, it was arranged for the child to remain in school with the exception of one hour a day, during which she was to be tutored by the writer.

In order to compare her progress with that of a normal child of her own age, a child was selected from her class who by the Stanford Revision of the Binet-Simon tests was found to have an IQ of 100 and whose mental and chronological ages were 6 years, 6 months. This child, Ray S., received only the class instruction. The study extended over a period of 4 months but only 3 months of actual work was accomplished.

Before beginning to tutor Grace, an effort was made to find out how much she had gained during her two years of schooling. She read from memory several pages from different primers. Stories which she had heard many times she repeated verbatim; of others, which she had

heard fewer times, she gave the thought correctly, but not as written. The primer pictures aided materially this memorization. When asked to give a certain word she began at the beginning of the sentence and read word by word until she reached the required one. When the first words of each sentence were covered she made good guesses but was sure of none of the words except "see." This one word she recognized wherever she saw it. In phonics she was hopelessly confused. She knew the sound of many of the letters but could not connect the right sound with the right letter. She could count by 1's and 5's to 50; read and write numbers through 10; recognize silver pieces of money; and discriminate fairly well size, length and weight. She copied writing quite accurately, but did not know what she was writing. She enjoyed the social subjects, language, story and dramatization, music and drawing, although her standing in all was below the average.

It was seen at once that teaching this child to read must be a process of teaching her individual words. Her past work showed that she was able to memorize the story sentence by sentence, and that a large amount of repetition and drill had not been sufficient to teach her one word from another. After she had gained a small reading vocabulary the following experiment was tried in order to see how rapidly she memorized. A new short story, containing twelve sentences and having no illustrations, was selected. As it was in the nature of a review exercise, the story was not particularly interesting, so the content did not give undue aid. The first time Grace read it she asked for 22 different words out of 67, the entire number. Every second or third day the selection was read once, and no attempt was made to teach the unknown words. During the fourth reading she asked for only 5 words, and at the sixth reading she knew it perfectly. I then covered parts of the selection and pointed out one at a time the original 22 unknown words. She recognized none of them.

She had been taught reading chiefly by means of visual and auditory associations. Proceeding on the basis that increasing the number or kind of associations would increase the possibility of recall, we planned to increase gradually the kind of associations made with each word and note the results. But in order to have a check upon possible future progress, during the first 3 weeks of the tutoring only the two forms of association, visual and auditory, were used in the following manner. A simple story was selected from the primer which Grace chose. After I made sure by means of pictures, discussion and dramatization that the meaning of the story was clear, we concentrated

on a small unit of it. This unit was read and re-read, the sentences read in order and out of order, and a few phrases and words found and matched.

In order to show the method and the time necessary to teach one word, a brief review of one lesson near the beginning of the study will be given here. The paragraph which follows is a small unit of a well illustrated story in which the mother cat is trying to find food for her kittens.

The cat saw a bird.
The kittens saw it too.
The bird saw the cat.
It saw the kittens too.
The bird flew away.

By reading the sentence she could name any word in the paragraph. With the first word of each sentence covered, she knew only those which we had previously studied. In this lesson she was learning "bird." She found the word as many times as she could in the paragraph, first with the entire paragraph exposed, then with the first word of each sentence covered. The word was written for her each time she found it. She then picked out the word from a small number of cards and matched it as many times as she could. This word was then put with several other review words and we played word-games for five minutes. At the end of this time she might or she might not recognize the word.

Her attitude toward the work was good. She wanted to learn, but her ability to concentrate varied greatly. Some days she would start to work with feverish concentration which lasted from 5 to 60 minutes. When a change of occupation was suggested she would beg to read "just the next page." On these days I could put all the words previously studied into a new story, the word arrangement entirely different, and she could read it easily. Never did she want to be told review words. Perhaps the next day she would be extremely flighty. Her attention could be held only by means of strong incentives and frequent change of occupation. On these days she remembered only the words on which we had drilled the longest length of time. This variability of concentration characterized all of her work throughout the 4 months.

In the first 3 weeks she learned 7 (!) very easy words, but in addition to this she began to understand the reading process. Formerly,

reading meant repeating something from memory or from imagination while she kept her eyes on the book. Now, she could not only differentiate between a few words, but she knew that she knew them.

It is generally recognized that a moderate amount of phonetic study improves the ability to read. However, if it is started too early or over-emphasized during the reading period it is apt to become an end in itself and block comprehension of thought.

In the case of Grace the phonetic training had been a hindrance rather than an aid in word recognition. She knew a number of the letter sounds, but with the exception of "S" could not associate the correct letter sound with the letter form. When asked to sound a word, as she was expected to do in her group work, she gave the first sound she thought of, regardless of letter form. Obviously, the thing to do was to discontinue all phonetic study until the point had been reached where it would be an aid rather than a source of confusion. But, because this was not possible in her group work, it seemed advisable to begin at the beginning and make the process as meaningful as possible, keeping it entirely separate from our reading period.

Consequently, each day we had a very short phonetic period beginning with games in slow pronunciation which were designed to help her hear correctly and recognize spoken sounds. Paralleling this oral work, the known sight word "run" was separated into its three sounds on the blackboard. She had little difficulty in locating and sounding the letters in order, but at the end of the three weeks she could not give the sound of "u" or "n" without beginning at the first of the word and sounding the letters in order. For instance, when "u" was shown her either on a separate card or in another word, she always went through this roundabout process: "It's in run, Run--r-u-n; *u*. Toward the end of the three weeks the known word "cat" was analyzed in the same way with the same results.

At the beginning of the fourth week, the following change was made in the reading period. Each day after the small unit of the story had been read, Grace chose one word for study. Then in addition to the visual and auditory associations, as described before, kinesthetic associations were emphasized. The word was pronounced and written for her several times on the board, while she watched. The work was then erased and she attempted to write the word. If she had difficulty, it was written for her again until she was able to reproduce it, but as her visual memory for form was quite good, this was not a tedious process. She then wrote the word as many times as she could in 5

minutes, or until her interest lagged, pronouncing it as she wrote it. The word was then, as before, put in the review list and used in games.

Although her ability varied greatly from day to day, it was evident that the addition of kinesthetic associations was beneficial. In case she could not recall a review word her response invariably was, "Don't tell me, I can get it if I write it." Although she sometimes wrote the word several times, if she did not recall it during the first or second trials, additional writings did not do so. When she could not recall it by writing I showed it to her in a familiar sentence; if this failed the word was pronounced for her.

During the 2 weeks in which the 3 types of association were made for every word, she studied 9 different words, 6 of which she learned comparatively easily and remembered. She was able to read by this time some 10 pages of a primer, the words of which she actually knew.

In phonics no new work was taken, but the kinesthetic association was made with each of the 6 letter sounds previously mentioned. At the beginning of each lesson it was necessary for her to go through the round-about process already described, then she remembered the sound until something else intervened. In the games, speed was gradually emphasized; but by the end of the fifth week, although she could give practically instantaneous responses to the four sounds, *r*, *n*, *c*, *t*, it seemed certain that the key word was inwardly vocalized before she was able to respond. Much improvement was noted in her ability to hear and give sounds correctly.

At the end of the first five weeks of tutoring, Grace was required to be out of school because of infectious skin trouble. Vacation came on and it was 4 weeks before her lessons could be resumed. During this time she was given no school help. She was glad to get back to school. Her teacher reported that she put forth more effort to learn and was easier to control than she had ever been.

I gave her a thorough review of everything she had studied with me, and found she had not forgotten. She read everything she knew before vacation and knew the individual words. She remembered every word in our word list with the exception of the last word studied.

During the rest of the time Grace was tutored (5 weeks) she received but 3 hours a week. Tactile associations were now added. The reading and phonic study continued in much the same manner as described. Short stories were chosen, for she soon tired of one and

wanted another. New stories were made on the blackboard using as far as possible her known vocabulary. She continued to study one word a day. The time spent on each word was gradually reduced to 4 or 5 minutes. A slow but steady progress could be seen. The combined associations, visual, auditory, muscular, and tactile added enjoyment and variety to the lesson which helped to hold attention and effort, and also reduced appreciably the length of time and the amount of mental stress necessary to learn a word.

The following table shows the ability of both Grace and Ray in reading and phonics; (*a*) when the study was begun, and (*b*) when it ended. The amount gained by Grace was found by averaging the results of thorough reviews given on 3 consecutive days.

When it was determined that 4 or 5 minutes was the least possible amount of time in which Grace could learn a word with a fair assurance that it would be remembered at least a few hours, I tried out Ray's ability. At 3 different times, 10 days apart, a list of words which he miscalled or did not know were taken from his reading. Five minutes directed study was given to each list which contained respectively 11, 12 and 14 words. Three days after each list was studied, I tested him and found that he remembered or was able to work out for himself 10 of the first list, all of the second, and 12 of the last. A little less than ½ minute was necessary to teach him a word.

Comparative Table to Show the Ability of the Two Children in Reading and Phonics When the Study was Begun and at the Close, Three Months Later.

ABILITY WHEN THE STUDY BEGAN

Reading		Phonics	
Grace	Ray	Grace	Ray
None.	Ray read with ease the hardest stories from three different first readers of average difficulty. He grasped the thought easily with one reading.	Knew one consonant, *s*.	Knew all consonants and short vowel sounds. Could sound new one- and two-syllable words containing short vowel sounds.

ABILITY WHEN THE STUDY ENDED

Reading		Phonics	
Grace	Ray	Grace	Ray
Has a reading vocabulary of 38 words. Reads the easiest parts of several primers. From 8 to 15 pages each.	Reads easily and with understanding in any of several second readers which he has not seen before.	Knows 13 sounds. Can sound one-syllable phonetic words which contain these sounds. Learns one sound or one word in a 5-minute directed study period.	Has learned the long vowel sounds and 21 phonograms. Can sound independently nearly any phonetic word found in a second reader of average difficulty. Learns 10 or 11 words in a 5-minute directed study period.

It will be seen by the comparative table that at the close of this study Grace had a reading vocabulary of 38 words. Three months after her tutoring stopped, I gave her a review of the work she had done with me and of the work she had done since that time. Of the 38 words she remembered 23. The forgotten words were mostly the unusual words which she had not had in her recent reading, such as *Christmas*, *dress*, etc. Her teacher was now giving her and another defective child a daily 10-minute reading period. They were using a very simple text and not attempting to keep up with the regular class. Under these conditions she was able to remember nearly $\frac{2}{3}$ of the words because she was using them daily and was not forced too rapidly into difficult material.

Grace's language ability exceeds her reading ability. She is interested in what is going on about her, and particularly in out-of-door things. Her guinea pigs at home, the pups in the neighbor's yard, a bird she had seen that morning, and the endless changes brought about by the seasons are all objects of her observation and interest which she talks about freely to anyone who will listen. Yet she displays little thoughtful curiosity. Her questions are not of the usual type which want to know "why" and "how;" she thinks emotionally rather than intellectually.

In the reproduction of easy, familiar stories Grace does as well as Ray. The interpretations of both are good. With less familiar stories Grace does very poorly. She is very fond of hearing stories. I found it advantageous to start the hour's work by reading to her, for she never failed to calm down and listen quietly as soon as the story started. I always questioned her closely to see if she grasped the thought and found that she did. The books read to her were those classified for 7- and 8-year-old children. When she was one of a group of listeners, her attention and interest varied.

The Stanford Revision of the Binet-Simon Tests gave the following vocabulary scores for both children: In the first test Grace made a score of 10, and Ray a score of 19, neither passing the 8-year-old requirement, which is a score of 20. Five months later Grace scored 20 and Ray 26. The re-test showed no change in Grace's IQ.

The Thorndike Reading Scale A-2, X series, was given to both children in oral form as neither could read the lists of words. The tests were scored by the method used in the 1917 Nassau County Survey; *i.e.*, counting a score of one for every word correctly classified. Ray scored 45 and Grace 39.

Graded in "usual quality" on the Thorndike Scale for Handwriting of children, writing samples of both Grace and Ray ranged between 9.3 and 9.9. The norm for grade II is 7; grade V, 9.3; grade VI, 9.9. In speed Grace averaged 19 letters a minute and Ray 21. The second grade norm is 35 letters a minute.

Grace enjoys the drawing and handwork period more than any other in the school program. She occasionally does a good piece of work; but unless her teacher stands by her and insists that she follow directions she cannot hold to one idea long enough to finish it.

While this study was in progress Grace's conduct showed some improvement. Realizing that her bad habits were partly due to her home conditions, Grace's teacher and I worked with the mother and older sister, trying to show them how we handled her at school. Grace daily reported to me her conduct at home, and her reports were checked up. After the first few weeks her improvement was quite marked. She knew she would be called upon to account for her acts at a definite time and that untruthful statements would be discovered. At the close of the study she was no longer the outstanding character in the room. When objects were missing, she was of course suspected, but in the last 2 months of the study she was found guilty only once, and in that instance she returned the article before it was discovered

in her possession. As far as we knew she had not destroyed other people's property. She had not thrown herself into a passion—a former frequent habit—since the Christmas vacation.

However, we do not expect improvement in this field to be of a permanent nature. Her conduct will always vary with the kind and amount of supervision exercised over her.

In conclusion, the following points were found to be most helpful in teaching this defective child.

1. Special instruction was necessary. She could learn nothing in a group of normal children except auditory memory work. As the group instruction became more advanced, her mental confusion increased.

2. The necessity of planning each lesson step by step and presenting in each lesson only a very small amount of new material was apparent.

3. The daily review had to be more detailed and thorough than with normal children. Words not used regularly were forgotten.

4. Perhaps the most significant part of the study relating to the method of instruction was the result of using the kinesthetic associations in teaching each word. The visual and auditory associations by themselves were not effective, but when they all were used, with the emphasis upon the kinesthetic, the word was taught more quickly and with more assurance that it would be remembered.

It is obvious that Grace should not be in a class with normal children. It is not fair to her, to the rest of the class, or to the teacher. She is now 9 years old. In the course of the next 2 years, unless a change is made, she will become very restless and dissatisfied when the looked-for promotions do not come.

To protect society from an increase of her kind, the only place for this child is in an institution for the feeble-minded, where adequate supervision is assured. She has no special ability to be trained, but she could be taught to do the type of thing she enjoys most; *viz.*, out-of-door work.

JUDGING HANDWRITING

A Critical Weakness of the Thorndike Scale Revealed by a Comparative Study[1]

WILLARD W. BEATTY

Presidio Open Air School, San Francisco, California

Since the Thorndike, Ayres and Starch Scales for Measuring Quality of Handwriting have come into general use, several comparisons of their relative efficiency have been made. Rudolf Pintner, in The Journal of Educational Psychology for November, 1914, published a study which he thought proved a slight superiority for the Thorndike Scale. The month following, Truman Kelley demonstrated that through an error in computation the conclusions were erroneous, and that the superiority, though slight, really lay with the Ayres Scale. Leroy W. Sackett in *School and Society*, October 26, 1916, published a comparative study also tending to favor slightly the Ayres Scale, and F. S. Breed in *The Elementary School Journal*, February, 1918, showed by a similar study that the Gettysburg Edition of the Ayres Scale yielded more accurate results than the earlier three-slant form.

The upshot of the whole discussion has been that all three of the above mentioned scales have been demonstrated to yield almost the same results in practice—to the extent that fairly accurate constants have been derived for translating values on one scale into comparative values on the other.

The data here presented were secured during a similar study, which yielded results indicating that little if any superiority could be claimed for any particular scale judging from comparison of deviations in judgment. The graphing of the results of the scoring does indicate a definite weakness in the Thorndike Scale, which may be of interest.

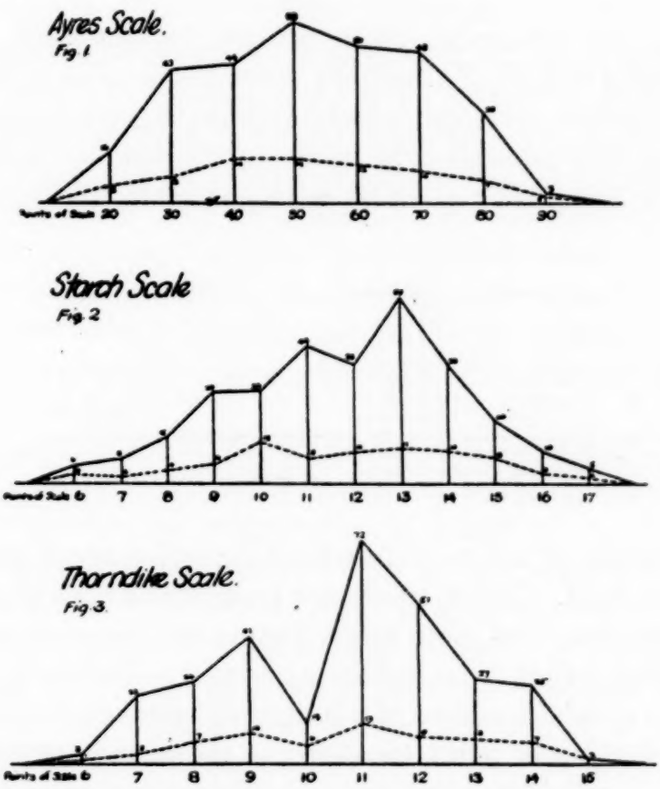
In the study, 17 typical handwriting samples were chosen from among 500. They were scored at intervals by 17 different persons, on each of three scales, the Ayres, Starch and Thorndike. After the remaining 483 papers were scored, 9 of the judges re-scored these 17 papers without knowing that they had been repeated. The results were tabulated, raised to a uniform scale for comparison and graphic charts prepared. As most of these data are very similar to those published by others, they are omitted. The method of graphing brings

[1]The data used in this article are drawn from a handwriting survey of the Daniel Webster School, San Francisco, carried out by a Seminar in Educational Measurements at the University of California, under Dr. Cyrus D. Mead, 1920.

out what would appear to be a fatal weakness of the Thorndike Scale, and is here presented.

Plate I shows the graphs. In all 289 judgments were made on each scale—17 papers, scored by each of 17 judges. These judgments were distributed over a certain number of points——that is, (Fig. 1)

Plate I



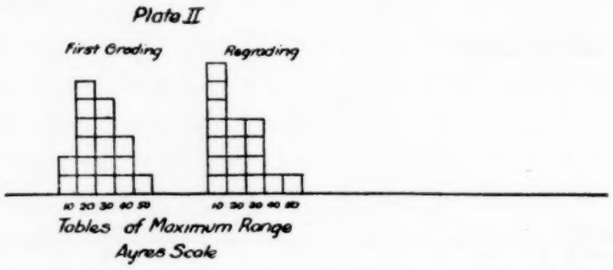
Curves Showing Frequency of Use of Points on the Ayres, Starch and Thorndike Scales
Solid line - Number of Judgments at each point (total - 289)
Dotted line - No. papers with one or more judgments at each point.

out of a total of 289 judgments, point 20 on the Ayres Scale was used just 16 times, point 30 used 42 times, etc., as shown by the black line. These 16 scores of "20" were divided among 5 different papers; the 42 scores of "30" were shared by 8 different papers, and so on, as shown by the dotted line.

If the papers represented typical samples over a wide range of the scale these two curves should be fairly uniform for all three scales. Referring to the graphs, it is seen that Fig. 1, showing the distribution for the Ayres Scale is fairly uniform. Fig. 2 for the Starch Scale is a

little ragged, which may be accounted for by the fact that Starch offers half again as many samples in about the same maximum range. Fig. 3, for the Thorndike Scale reveals a remarkable drop at point 10, the judgments falling from a total of 41 at point 9, and 72 at point 11, to only 13 at point 10. The scores which might have been expected to fall at point 10 are divided between 9 and 11—11 getting the lion's share.

A glance at the Thorndike Scale reveals the difficulty. Only one sample of writing is used to illustrate point 10 as against two or three at every other point; and most unfortunately, this particular sample is a very peculiar and unusual type of almost back-hand writing. It is natural that in comparing samples with the scale, the judges should



seek an illustration of a type of handwriting which most nearly resembles the sample held. Point 10 cannot possibly be typical of 1 sample in 1000. In the long run, with many judges scoring each sample, and an average being struck, this deficiency would probably be overcome, but in any survey where one or two judges did most of the scoring, this unfortunate selection of point 10 added to the general unwieldiness of the scale, should eliminate the Thorndike Scale from consideration.

Plate II shows the result of practice in scoring with the Ayres Scale. The first diagram represents the maximum range on each of the 17 samples when first scored, the average of which will be seen to be about 27 points. The second diagram shows the change in maximum range on the re-grading, the average of which is a little over 21 points. It will be seen from the diagram that the number of papers where the range was over two points on the scale has risen from *two* on the original scoring to 7 on the re-grading—a "coming together" on the part of the judges which is quite remarkable.

Extended use of the scales, as well as the data presented, served to convince the judges that for convenience, reliability and all-round practicality, the Ayres Scale (Gettysburg Edition) is to be preferred for general use.

THE TORONTO MEETING OF THE SECTIONS ON PSYCHOLOGY AND EDUCATION OF THE AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

FRANK N. FREEMAN

University of Chicago

The recent meeting of the American Association for the Advancement of Science was the second one to be held in Toronto. In addition to a good attendance of representatives of psychology and education from the United States, there was a considerable attendance of Canadian psychologists and educators. The association between the workers from the two countries was very profitable, at least from the point of view of the visitors, and opened their eyes to the amount of scientific work which is being carried on in Canada. Some of this work was represented in papers which were read at the meeting by the Canadian scientists.

In the first session of Section I, two general papers were presented by men in the University of Toronto. J. A. Dale discussed the place of psychology in university curricula, and in his discussion represented particularly the point of view of the practical social sciences and more generally the demands of the applied sciences. He stressed the desirability of treating psychology in such a way that it would meet the needs of the groups engaged in practical enterprises. G. S. Brett discussed in a discriminating manner the various schools or movements, such as behaviorism, which have been represented in psychology of the twentieth century. W. B. Pillsbury also presented a paper of general nature. He brought the various points of view in modern psychology into opposition and raised the question whether these opposed points of view might be reconciled. He inclined to the belief that they represent permanent and perhaps temperamental differences, but that they are irrelevant from the point of view of progress of the science. A. P. Weiss discussed the fact of decreasing variability with increasing age, and speculated upon the possibility that society may prolong somewhat this period of plasticity.

As in recent years, many of the papers were devoted to mental tests. The results of testing children upon entering school and the divergence in ability in children of the same chronological age were reported upon by L. W. Cole. Several papers dealt with the technique of tests or with general problems relating to their value or particular

significance. The standardization of tests was discussed by Peter Sandiford. He emphasized particularly the necessity of restandardizing tests for use in Canada which are imported from the United States. Methods by which the validity of intelligence tests could be determined, largely on the basis of correlation, were discussed and illustrated by Raymond Franzen. R. M. Yerkes argued the necessity of considering psychological examination to be a much broader thing than intelligence testing, and maintained that it is necessary to use a more refined technique for research upon mental capacities than for application of intelligence tests.

A critical discussion of the application of intelligence tests to college students was presented by J. W. Bridges, who maintained that the prevailing correlations were so low that tests are of much less use in college for classification or other administrative use than is popularly thought. On the other hand, papers by G. M. Whipple and Wilhelmina Koerth described the application of intelligence tests to students in the universities of Michigan and Iowa. Whipple applied the tests to students on probation and concluded that other factors than intellectual ability were frequently the chief causes of failure. He therefore regarded lack of correlation as not necessarily evidence of the unreliability of the test. Miss Koerth showed that there was a wide divergence in the character of work done by those who stood in the upper and lower tenths in the intelligence test. A paper on the educational significance of mental tests, by William D. Tait, maintained that their results are to be accepted as indicating that education should be much more selective than it now is since there are many individuals who are incapable of profiting by it.

A number of papers on miscellaneous subjects may be mentioned. A paper on the psychology of the equation by E. L. Thorndike drew the distinction between two types of equation, the one being a mathematical statement which is made for the purpose of arriving at a solution, and the other being a statement of a relationship between a variable and one or more other variables. The first represents the common use of the equation in the conventional algebra, and the second represents its use in formulae, such as those which define certain types of graphic curves. The use of these two types is very apt to lead to confusion, and the author of the paper suggested various means by which this confusion might be avoided.

A paper by B. T. Baldwin and Lorle I. Steckher presented graphs showing the results of repeated mental tests and of repeated physical

tests of average and superior children. The results suggested that the mental growth curves of superior children diverge from those of average children, and that mental curves resemble and run parallel to physical curves. The inadequacy of children's concepts regarding matters which are assumed in their instruction was described and illustrated in a paper by Garry C. Myers. It was reported, for example, that many children who could recite the fact that Franklin was minister to France thought that this meant that he was a clergyman. A valuable study of the inmates of the Illinois penitentiaries was contributed by Hermann N. Adler. This study agreed with a recent study in Ohio and in general with the results of the tests in the army in showing that the inmates of penal institutions do not differ in general intelligence from the average of the population. The Army Alpha Test was used in these examinations. It was reported, however, that these prisoners in large numbers exhibited anomalies of behavior or of mental attitude. These anomalies, however, are not usually so serious as to make normal adjustment to society impossible. The somewhat related subject of psychiatry in the public schools was presented by Eric K. Clarke with illustrations from his experience in Toronto. T. R. Garth reported the results of a study which showed that pure blood Indians have different color preference from Whites.

One session was devoted to the applications of psychology, particularly in industry. The functions of an industrial relations department, with certain comments on the psychological aspects of the activities of such a department was presented by George W. Allen. The difficulties and psychological problems of job analysis were discussed by E. K. Strong, Jr. He pointed out that job analysis is the determination of the habits that a man must have to perform the work of the job and of the native qualifications that are necessary to enable the man to acquire these habits in a reasonable length of time. In the higher positions, however, the job is something more than the sum of its parts. A descriptive account of problems which confront the handicapped in securing and retaining work was given by Norman L. Burnett. He emphasized the factors of general attitude and personality in judging the prospects of success of a handicapped man and in placing him in a position. Alfred E. Lavell presented the successful experiment which is being made in Ontario to bring about a satisfactory attitude on the part of prisoners by means of employment outside the prison walls.

The writer was not able to attend two of the sessions of Section Q and therefore cannot report upon them.

The vice-presidential address before the section on psychology delivered by Dr. Strong, dealt with Control of Propaganda as a Psychological Problem. A psychological analysis of propaganda was made and supported by numerous illustrations. The paper emphasized the fact that propagandists influence the public by arousing sentiments and connecting these sentiments with particular forms of expression through suggestion. Our present methods of legal control are not adapted to this form of propaganda. Publicity does not seem to be a sufficient safeguard and an adequate solution of the problem is not yet at hand.

Dr. Judd, Vice-president of the section on Education, dealt in his address with the problem of the use of the scientific method in constructing the curriculum. He showed that there is at present no agency whose business it is to collect curriculum materials or to subject them to adequate standardization and drew illustrations from a recent attempt to develop and organize material for courses in the social studies. He offered no definite solution of the problem but presented it as one worthy of serious study.

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES

REPORTED BY CECILE COLLOTON

Department of Educational Psychology, The Lincoln School of Teachers College

Intelligence and Educational Tests

What Los Angeles is Doing with the Results of Testing. Harlan C. Hines. Journal of Educational Research, 1922, January, 45–57. The use of the "intelligence survey" in organizing Ungraded and Adjustment Rooms in the elementary grades. Discussion of plans for a new marking system and the use of mental ability scores in educational and vocational guidance in the secondary schools.

Intelligence Tests in the Primary Grades. M. Edith Whitcomb. Journal of Educational Research, 1922, January, 58–61. The use of the Stanford-Binet in the grading and promoting of 2,360 primary children in Council Bluffs, Iowa. Comparison (in per cents) of quality of work and IQ's.

Intelligence Tests in Massachusetts Normal Schools. E. A. Kirkpatrick. School and Society, 1922, Jan. 14, 55–60. Results of administering the Thurstone Test to all students in the normal schools of Massachusetts. Comparison with scores of high school and college students.

Miscellaneous

A Diagnostic and Remedial Activity in Supervision. Bertha M. Rogers and Teresa Baker. Journal of Educational Research, 1922, January, 21–26. The use of the Woody Tests to show the need of improved instruction. Details of the experiment. One illustrative lesson used in remedial work is given in full.

The Correlation of Visual Memory and Perception of Perspective with Drawing Ability. Elmer Jones. School and Society, 1922, Feb. 11, 174–176. Preliminary report on an investigation being carried on at Northwestern University to determine native powers peculiar to children possessing art ability. Description of two tests for the special traits of visual memory and perception of perspective.

Problems and Solutions in Classification of School Children. Jennie B. Boyer. Detroit Journal of Education, 1921, November, 30–33. A general discussion covering retardation and its causes, individual differences, and the advantages of individual instruction.

Some Problems Arising in the Administration of a Department of Measurements. Helen Davis. Journal of Educational Research, 1922, January, 1–20. Discussion of problems such as Cooperation with Teachers, Test Administration, Classification, Publicity, etc., likely to be met by a director of measurements. Suggestions for principles and methods of solving such problems.

Provisions for Individual Differences in High School Organization and Administration. W. H. Hughes. Journal of Educational Research, 1922, January, 62–71. Are any provisions made in your school for individual differences of pupils? Summary of answers received from 221 high schools with an average enrollment of approximately 1000 each. Forty items are included in the tabulation; *e.g.*, basis of classification, supervised study, social recognition for superior students, etc.

The Basis of Individual Differences. Alfred Hall Quest. Detroit Journal of Education, 1921, November, 8. First section of an outline used as the basis of a series of lectures in the Detroit Teachers College. The outline will be continued in later issues.

Mind Set and Learning. Wm. H. Kilpatrick. Journal of Educational Method, 1921, December, 144–150. Continued from the November number. A popular presentation of the laws of learning.

Material in Four Delayed Issues of the Journal of Educational Psychology

Two Important Points with Regard to Age-grade Tables. S. L. Pressey. Journal of Educational Psychology, 1920, September, 355–360. A suggestion for the use of median age per grade and percentiles instead of per cent of retardation and acceleration, or age norms.

Superior Children—Their School Progress. Anna Gillingham. Journal of Educational Psychology, 1920, September, 327–347. Studies of gifted children in the Ethical Culture School and their achievement in school subjects. Descriptions of individual cases. Particular needs of the gifted child.

Solution of Problems in Geometry. Ben Wood and J. Carleton Bell. Journal of Educational Psychology, 1920, September, 316–326. An experiment to determine some of the chief factors in geometric ability. Descriptions of tests used and details of results.

Tests for Mental Alertness. L. W. Sackett. Journal of Educational Psychology, 1920, November, 430–444. Description of a test consisting of a short story and a set of exercises based on the story. Directions for administering, scoring, and interpreting results. Detailed tables.

Intelligence Ratings by Group Scales and by the Stanford-Binet Revision of the Binet Tests. G. M. Ruch and Lexie Strachan. Journal of Educational Psychology, 1920, November, 421–429. Correlation between scores on Army Alpha and Mental Age as determined by the Stanford Revision of the Binet-Simon Scale = 0.728; between Chicago and Stanford-Binet $r = 0.622$. Other correlations given with separate parts of the tests.

Tests for the Measurement of Certain Phases of Linguistic Organization in Sentences. Harry A. Greene. Journal of Educational Psychology, 1920, December, 517–525. Description of a language test for sentence organization.

A Comparison of Results Obtained by the Terman-Binet Tests and the Healy Picture Completion Test. E. B. Skaggs. Journal of Educational Psychology, 1920, October, 418–420. Extremely different ratings resulting from the Terman Binet and Healy Picture Completion Tests. Correlation = 0.15 ± 0.08.

Further Data on the Bell Chemistry Test. H. L. Gerry. Journal of Educational Psychology, 1920, October, 398–401. Results of the Bell Chemistry Tests in Biddeford, Maine, High School and Worcester Academy. Original standards too low.

The Correlation between College Grades and the Alpha Intelligence Tests. J. W. Bridges. Journal of Educational Psychology, 1920, October, 361–367. Results of the comparison of college grades and standing on the army tests of 486 students at Ohio State University. Four tables of interesting correlations.

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION

Books on New Mental Tests

Mental and Educational Tests in England.—Two recent contributons,[1] in the field of psychological testing show how steadily the movement is advancing in England. Mr. Burt's treatise is the most considerable contribution to the literature of the subject that has yet appeared in that country, and the preface is over-modest in setting forth its aim as that of presenting "a provisional set of practical scales for measuring intellectual ability and educational attainments" in the elementary schools. It not only does this, but also presents a vast amount of useful critical and statistical material. The scope of Mr. Godfrey Thomson's work is much narrower. It is an account of the group test which he devised in order to discover gifted children in localities of small educational opportunity.

The first section of Mr. Burt's book consists of an English version of the Binet-Simon scale, which was begun with the collaboration of M. Simon, and in its final version is the result of 8 years of experiment with over 3,500 children. The directions for giving and scoring each test are extremely careful, and there is an excellent key by which a test-score can be converted into a mental age in a few seconds. In discussing the various influences affecting proficiency in the tests, —schooling, sex, social status, etc.— Mr. Burt emphasizes the importance of schooling, and though the reciprocal relation between intelligence and educational attainments receives attention, he finally uses a partial-correlation equation to prove a proposition which will possibly be challenged, namely, that in determining the child's performance in the Binet-Simon scale, "intelligence can bestow but little more than half the share of school." In estimating the value of the scale for practical purposes he considers it greatest for use with defective children, less with normal children, and least of all for detecting the specially

[1] Burt, Cyril: "Mental and Scholastic Tests." London, P. S. King & Son, 1921, pp. XV + 432.

Thomson, Godfrey H.: The Northumberland Mental Tests. *British Journal of Psychology*, December, 1921.

gifted child. "The unwarranted claims advanced on its behalf by votaries in foreign quarters have among academic psychologists in this country become a commonplace and a byword." Yet as a provisional, workable substitute for a scientifically exact scale, he urges its use. His reason for standardizing for England the original Binet-Simon scale, instead of making a more or less drastic revision, as has been done in America and Italy, is that a present, uniformity can be secured and confusion avoided, only by this means. He hopes for a future English revision "which will ruthlessly abandon the tests which are now known to be worthless—sex, surname, date, months, two lines, comparing faces,—"and will consistently pursue the principle that has been fitfully attempted in America, the "internal grading" of each test, so that it will appear at the different age-levels with regular increments of difficulty—the method which Terman has adopted in the case of his vocabulary test, and that of repeating digits backwards. As a useful rough method for testing new tests, he suggests the use of Yule's colligation coefficient ω, the Coefficient of Association; an Appendix describes this statistical device in detail.

Besides the Binet scale, Mr. Burt gives one form of each of the usual intelligence tests, standardized with London children, Synonyms, Definitions, Instructions, Completion, Absurdities, Maze, and as an example of the thoroughness of his methods, we may point to the 16 criteria he employed in compiling his list of opposites (p. 223). It is interesting to remember that Mr. Burt was the first psychologist to use what is commonly known as the "Analogies" test to measure intelligence, when he was working in Liverpool in 1911, though Woodworth had analysed the mental processes involved in the recognition of such relationships 3 years earlier. Mr. Burt gave the test its name from Aristotle's ἀναλογια, proportion.

The intelligence test which he considers the most efficient, especially for older and brighter children, is his individual test of Reasoning Ability, which consists of 17 brief questions, preceded by the data necessary for answering them by ordinary logical inference. Each problem is assigned to the age at which 50 per cent of the children answered it satisfactorily. This test was used successfully last year in a small experiment in a Public School in Brooklyn.

The section of the book also gives some interesting graphs of the Distribution of Intelligence among London children, both in the ordinary elementary schools and the "Special" schools for Mental Defectives. After much discussion of this highly debatable point,

Mr. Burt suggests a Mental Ratio (IQ) of 67–70 as a theoretical line of demarcation for mental deficiency in children, *i.e.* to justify removal to a Special M.D. School. As provisional borderlines at the upper end of the scale he suggests that a Mental Ratio above 115 (or even 120) denotes central school ability, (Central Schools are lower secondary schools, usually having a commercial or industrial bias,) and that one above 130 (or 135) merits a full secondary school course. If this last estimate really represents the lowest level of intelligence of the "free-placers" in secondary schools, it shows the futility of comparing the average intellectual attainment of an American High School with that of a London Secondary School. In no public elementary school did Mr. Burt find an IQ higher then 160, though in a private school he tested a boy of 7 with one of 190, as high as that of the recently reported American and Scotch precocities.

The last section of the book is devoted to Educational Tests and Scales for the elementary subjects, with a description of individual cases of disability in each. The author gives norms for each age, but adds a warning against the danger that lurks in thus giving prominence to the average or median score. Just as an official minimum wage tends to become the maximum, or at least to limit it, so a risk arises lest better performances should tend to be depressed towards the mean. He would be startled to hear the devotees of standardization acclaiming this as an advantage. The Drawing scale consists of a median specimen of a drawing of a man for each age from 3 to 14. One wonders whether it is only a curious coincidence that in 5 of the specimens the man is smoking!

The whole book gives evidence of the scholarship of the author, and leaves one with a strong impression of his intense interest in the personality of the individual child.

For nearly a quarter of a century, poverty, unless very dire, has not excluded really gifted children from full educational opportunity in England, provided that they lived in or near a town. The ideal has been the "ladder" whose rungs are scholarships, sometimes with maintenance grants, leading from the Public Elementary School through Secondary Schools of various types, and perhaps a Provincial University, to the best that Oxford and Cambridge can offer. The scholarships to secondary schools are given to children aged about 11, who are presented for examination in Arithmetic and composition by the elementary schools, and it has recently been noticed by one Education Authority, the County of Northumberland, on the Scottish

Border, that more than one quarter of its schools present no candidates at all, and that these schools are chiefly found in the remoter country districts. It was felt that two reforms were necessary to help the gifted children in these districts, (1) full maintenance must be offered in addition to free schooling, (2) the examination must be in the form of a mental test which would give an equal or almost an equal chance to an unprepared child. Mr. Thomson undertook the task of preparing such an examination, and after considerable preliminary experiment he chose the following group tests: New Tests—(1) Hindustani Test, (2) Extra Number Test—Partly New Tests, (3) Middle Word Test, (4) Schema Test, (both based on the work of Prof. Stern of Hamburg, but new in the form of presentation) Old Tests, (5) Extra Word Test, (6) Number Series Test.

In choosing a large proportion of new tests, he says he was actuated in part by a desire to make coaching for the tests almost impossible, and in part by a desire to make the experiment as original a contribution as possible, and to increase the supply of mental tests. Much reliance has been placed on some of the tests he definitely rejected, *e.g.* the Completion test, but the common use of this as a teaching device in English schools would bar its inclusion in this particular series.

The new tests are interesting, particularly the Hindustani test, which may prove of special diagnostic value. Each is given in full in the article.

The tests were given to nearly 3000 children in addition to the scholarship candidates whom the rural schools sent in, and age-norms were tentatively fixed from the results. The scores seem to correlate satisfactorily with the Binet testing of the same children, which is still in progress. The highest IQ found by the new tests (174), was that of an 8-year-old boy in a small village in the heart of the Cheviot Hills; and a comparison of the distribution of intelligence in schools in different localities showed that the most intelligent children are to be found in the small country schools in the hills, the results there being even better than those in a well-to-do town suburb. Experiments in other parts of England have shown that children in large cities are on the average a year in advance of those in the country. But in this investigation, the only very large town, Newcastle, was omitted, as it has its own Education Authority. Moreover the Cheviot children are to a large extent the descendants not so much of farm laborers as of Border troopers. Mr. Thomson draws the tentative conclusion

that the highest ability is to be found close to cities and far away from cities, the intermediate areas being drained by selection. It will be interesting to discover whether this holds good for other countries too.

EDITH NEWCOMB
Teachers College.

---

3. *A Further Revision and Extension of the Binet-Simon Scale.*[1]—This manual presents a much revised and extended Binet-Simon scale. Its aim is to make it "possible for a teacher, a student of psychology, a parent, a worker in a juvenile court or in an institution for defectives or delinquents, to use the Kuhlmann scale readily and accurately after the reading of this volume."

The revision is based on the results of seven years of continuous work with about 7000 children and adults in the Minnesota School for the Feeble Minded, and in the public schools of Minnesota.

The manual consists of five chapters and an appendix. Chapter 1 describes the nature of the revision and extension of the scale including a brief discussion of the general accuracy of the method and the significance of the results.

Chapter 2 is a detailed discussion of the general principles of the year scale based on the three topics (1) the requirements of the individual test; (2) the construction of a year scale; and (3) the establishment of norms.

Chapter 3 gives complete information as to the best way to conduct an examination, how to use the scale in abbreviated form, and how to determine the mental age.

Chapter 4 describes the materials needed for the examination and presents detailed directions for administering the tests and scoring the responses.

Chapter 5 summarizes various comments on the individual tests.

The appendix consists of a table of intelligence quotients which "gives all quotients from 25 to 150 for the ages of three years to maturity and the mental ages of 3 to 15 inclusive." Dr. Kuhlmann says his tests show the maximum mental age to be 15 years. "Maturity" evidently means 15 years in chronological age also for the table of chronological ages stops at 15 years 0 months.

Of the various revisions of the original Binet-Simon Scale probably

---

[1] Kuhlman, F.: "A Handbook of Mental Tests." Warwick & York, Inc., Baltimore, 1922, p. 208.

the best known is Terman's, which is called the Stanford Revision of the Binet-Simon Scale, or more popularly, the Stanford-Binet. The outstanding differences between the Terman and Kuhlmann revisions are the increase in the number of tests from six to eight in each age group above two years, the extension of the scale at the lower end (Kuhlmann's scale begins with tests for the age of three months) and the use of a different method of scoring when the subject passes several tests in the highest age group of the scale. The same tests are given in years 13, 14 and 15 but they are scored differently for the different ages.

The aim of the author has been to give a larger place to actual performance as compared with verbal response and to insure a higher degree of accuracy in scoring the results of an examination. Therefore, practically all of the new tests consist of things to do such as counting dots in a square; tapping blocks in irregular order; spelling familiar words backwards; crossing out q, r, s, t, in a pied test; following directions in a confusing text; drawing upright forms in inverted positions. They are scored in terms of the time it takes to do the test and the number of errors that are made.

Examiners will be interested in using this scale and comparing its results with those secured from previous revisions. It is generally conceded that most of the revisions do not adequately measure mental age at the upper end of the scale and if this new scale proves the claims of its author in this regard it will be of distinct service.

CECILE COLLOTON
The Lincoln School of Teachers College.

---

4. *A Manual for Case Study* which is extremely compact yet definite and complete has been recently published by the California Bureau of Juvenile Research.[1] Comprehensive instructions, with sample data, are given for taking case histories of intelligence, temperament, physical condition, moral character, conduct, amusements, education, home conditions, etc., etc. Chapters are devoted to the scope and meaning of social case investigation, methods (such as interviews, correspondence, inspection of records, etc.), methods of evaluating data, the use of charts, symbols, etc. It contains tables or norms of height, weight, complete sample histories and excellent lists

[1] Williams, J. Harold and others: Whittier Social Case History Manual, Whittier, Cal. California Bureau of Juvenile Research. *Bulletin* No. 10, December, 1921, p. 98.

of references on each topic discussed. This booklet, which can be purchased for a quarter, merits use in normal school and college classes and could be studied profitably by educational workers in general.

A. I. G.

---

5. *Psychology Coming into Its Own in the Making of Children's Books.*—If you were stranded with no books for children to read from and had to write your own, how would you go about it? Very likely you would let the children write the stories which would make up their reading books. You would put into the books the things to which children of the different ages give spontaneous attention. You would like to be with the children throughout their waking hours, constantly tabulating and interpreting what they did and said. You would answer their requests for stories by telling them some, trying first one kind and then another; and living close to the children, studying and recording what they like and demanded you would not go far wrong.

You would find, for example, that 3-year-olds simply revelled in the mere recital in story form of the, to them, exciting things of the day, but to us the humdrum routine of life. When they, themselves, are in the center of the stage they are consumingly interested in the daily tasks of getting up in the morning, dressing, eating, playing throughout the day and the like. And if you did write books for children this way and then studied critically the books which children now read in the primary schools, you would be impressed by the fact that someone had sold the primary pedagogues a good sized gold brick; for the said pedagogues have resorted thoroughly to Mother Goose, fairy stories, repetitional animal and similar stories.

Now, this very sort of thing is exactly what Mrs. Lucy Sprague Mitchell and her colleagues have learned by living with the children in the City and Country School of New York City. Mrs. Mitchell's *Here and Now Story Book*[1] is the result. Really, this book is quite revolutionary. No doubt, it will fearfully upset our writers of Primers and School Readers, 50-foot shelves of which are still being published each year. If they will only read Mrs. Mitchell's introduction, they will at least dimly see the why and wherefore of all the change, and then, if they will trail a little group of children for a few days—one would be quite enough—and concentrate very hard upon what the children really did and said, and use a good pinch of Woodworth's,

[1] E. P. Dutton & Co., New York, 1921, p. 360.

Thorndike's and Dewey's Psychology on it all—why they would, this reviewer humbly believes, go and do likewise. No teacher of primary children should fail to read the introduction to this book. And thousands of homes should turn to the stories in it to read to children: The Dinner Hour, The Grocery Man; The Journey; How the Engine Learned the Knowing Song; The Fog Boat Story; Hammer and Saw and Plane; The Skyscraper; Things that Loved the Lake, etc.

H. O. R.

A New Hand Book of Modern Education for Teachers.—In making a curriculum for school children one constantly needs to know the evidence which has been collected on mooted points of social needs, what children learn, what their interests are, what we know about their heredity, their intelligence and their growth. One of the essential tools of the progressive and experimentalist in school practice is an up-to-date, complete and described bibliography of what biology, psychology and sociology has to say about the nature of the child and the educative process. It would be still more helpful if liberal quotations were supplied of what our leaders of prestige have to say about these matters. Finally, the teacher needs definite suggestions concerning books and materials which she can use with children.

If I were a teacher with these needs I would get Miss Gertrude Hartman's new book *The Child and His School.*[1] It does those things and very well, indeed. *First*, it gives the teacher a good bibliography of books and materials (850 of them) that teachers can select from to be read by children in the lower and intermediate grades. These references are classified to fit: (1) Community study (food, shelter, clothing, transportation, communication, conservation of wealth, education, recreation, religion, government and primitive life); (II) National life (in general, its government, its history) ; (III)The study of other nations. True, many of the books listed are very poor readings for children, but, as we who are collecting and organizing reading materials know, they are all we have. This is an excellent list to have at hand.

Second, it does summarize, quote from and interpret for the teacher the scientific basis of education. This book is an example of what the Bureau of Educational Experiments (of New York City) is doing. It combines the use of a broad philosophic interest in child life and the

[1] E. P. Dutton & Co., New York, 1922, pp. XI + 248.

improvement of society with the scientific foundations of these in biology, psychology and sociology. It does not turn up its nose at the notion of measuring the child's ability to spell or that he must master certain arithmetical tools that will be needed in life. Most free educational institutions have displayed this attitude in recent years. Paralleling this characteristic of the Bureau to keep pace with the development of scientific education, they are committed to the working out (in schools) of the basic philosophy of the "free educationists" laid out with such clear vision by John Dewey. And Miss Hartman has made a very good illustration of this point of view in her book. She has also combined with it a keen use of scientific education.

H. O. R.

A New Statistical Book for Research Workers.—The scientific study of education waited on two things: *one*, the development of methods and devices of measurement; *two*, the taking over from biometry and mathematical statistics of the technique of statistical treatment of data. The publication of Professor Thorndike's *Mental and Social Measurements* in 1903 was epoch-marking. From that date on, for ten years, hundreds of technical publications in education appeared and rarely indeed, it was that footnote references to his book did not abound in the publications. My own *Statistical Methods Applied to Education* (1917) doubtless served to interpret on an elementary plane, methods by which the school man could treat his data. Neither of these two books, however, supplies the mathematical foundation nor the refined methods which research students in psychology and education need for the extension of their very scientific work. American students have been forced to used the elaborate publications of Karl Pearson, Yule, Elderton, Edgeworth and others. Professor Jones' new book,[1] especially part II, will help fill this gap admirably. Part I (using economic and biological illustrations) gives the mathematical basis of much of what is already available in American books on measurement, variables and frequency distribution; classification and tabulation of materials; averages, dispersion, graphs and correlation. American students will regret that Professor Jones has not discussed partial correlation and that they will still be forced to go to Yule's detailed original publication. In Part II students will have

[1] "A First Course in Statistics." G. Bell & Sons, London, England, 1921, pp. IX + 286.

access to a treatment of probability, sampling, curve fitting and type distribution curves that will be very helpful, indeed. There is little doubt but that we are entering upon the second stage in the use of statistics by students of education. The "statistics is arithmetic" stage in which we used frequency, averages, dispersion and correlation, blindly, is practically passed. The more high-brow stage of using statistics for the discovery of scientific law and of prediction we are now just entering upon. The work of Kelly, Otis, Ruml, Rosenow and others provides important applications in the determination of the reliability of tests and in the study of "law" in learning, growth and the like. Schools of education will need at hand, therefore, in increasing proportion such technical tools as Professor Jones' book provides.

H. O. R.

Another Manual on the Technique of Teaching.—Each of the new books which I have brought together for review for this month makes an addition to our method of teaching—some very large, some small. The next book to be discussed can hardly be said to do so. Davis' *Technique of Teaching*[1] is but another of a long stream of such books written largely out of the experience and observation of the author and only partially in touch with the newest scientific and philosophic developments. It contains a chapter of general principles of teaching, followed by a series of chapters, one devoted to the teaching of each of the principal school subjects—spelling, reading and literature, composition and grammar, arithmetic, history and geography. The best thing in the book is the set of exercises at the close of each chapter. Beyond this I find little to justify the publication of another general book in this field. In dealing with the separate subjects little or no reference is made to the psychological facts now available from the results of measurement. To one's astonishment he finds reference in the bibliographies at the end of the chapters to almost none of the scientific analyses of learning in the different "subjects." Instead the author refers prospective teachers to other books in the same category as his own—empirical and *a priori* pedagogies. The book reflects acquaintance with pedagogical theories of 1910, but not with the virile dynamic psychology and the scientific method of attacking educational problems which are developing so fast in 1922.

H. O. R.

---

[1] Macmillan, New York, 1922, pp. VIII + 346.

9. *A Novel Approach to Reading.*[1]—Recent contributions to the literature of method reflect the unsettled state of educational thinking. One hesitates to be duly critical of constructive contributions because of the good they may do. But if, by over- or under-emphasis, they tend to obscure real values and create new problems, criticism is urgent.

Miss Watkins presents a method by means of which beginners may be taught silent reading. She puts that method forth succinctly and systematically. She makes the materials available so that anyone may attempt to do what she has found possible by actual trial.

We quote from Chapter I:

"Using the Silent Reading Method, each pupil in a class of forty, long before the end of the first school year, carried out without hesitation the following printed commands in the presence of the class and visitors:

'Tell that man sitting by the window that the spinning wheel over in the corner is older than the telephone, electric light, railroads and the United States.'

'Draw an oblong, a circle and a square. Put your age under the circle, then go and turn on the electric light that is over my table.'

'The man standing by the door and wearing a gray suit lives in Omaha. Go to him, shake hands with him, tell him your name and your father's name; then show by the number of your swings on the swing just how old you are.'

'Ask the lady with the blue dress to whisper to you the time she came to Iowa City, then show us the time on the large dial. Then ask her if she came from Denver on the Rock Island Railroad. Tell us what she says.'"

Miss Watkins then asks and answers the following question:

"How can an immature and entirely untrained mind be taught in the short space of the first school year to grasp fully and execute commands involving a knowledge of every part of speech and of so varied an assortment of actions?

"The result is achieved by treating the child's intelligence as a full grown intelligence which simply requires to be informed, in a logical manner and without interposition of obstructive methods.

"There is apparently no limit to the capacity for learning which the child's mind possesses, save the limitation of time and circumstances. There is no reason why a child should not learn anything valuable which it is capable of learning in as short a time as possible."

Surely Miss Watkins would not have us select materials, methods and vocabulary on so arbitrary a basis! Should the first grade child's reading vocabulary include such terms as "personal history," "hibernate," "salutation," "doilies," "abdomen," "objects," and "pro-

[1] Watkins, Emma: "How to Teach Silent Reading to Beginners." School Project Series. J. B. Lippincott Company, Philadelphia and London, 1922, p. 133.

ject?" Some of these words are used so infrequently that they are not found among the 10,000 words listed in the Thorndike word counts. Is it feasible to introduce new word symbols in categories? Is an action response sufficient guarantee of the meaningfulness of a symbol or is it possible that the symbol acts as a direct cue to a particular response, regardless of real meaning? Is there a danger of word sophistication and reading facility without sufficient background in experience? What does the following sign mean to a 6-year-old? "Please do not ask for credit." Does the child who responds to the flash-card "vulgar," by saying its opposite, have any clear notion of what it is all about?

The directness of the method has much to recommend it. The use of flash-cards is no doubt conducive to concentration. The action response is valuable. Children *do* learn by noting differences.

The method is highly stimulating. The boy who was "so proud of his work and achievement that he took his tightly rolled test to bed with him" worries us. The satisfactions are not those which usually engross six-year-olds. The psychology of the method is fundamentally at fault in another respect. Miss Watkins speaks of the child's mind as though it were a receptacle. Another conception of learning makes for a more critical evaluation of motives and procedures.

Fortunately we read Dr. Ernest Horn's introduction in time to realize that the proposed method is not supposed to displace "instruction in literary appreciation" and extensive reading from printed materials.

L. Z.